FINITE PERTURBATION ANALYSIS METHODS FOR OPTIMIZATION OF INVENTORY SYSTEMS WITH NON-STATIONARY MARKOV-MODULATED DEMAND AND PARTIAL INFORMATION

A THESIS SUBMITTED TO

THE GRADUATE SCHOOL OF ENGINEERING AND SCIENCE

OF BILKENT UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

THE DEGREE OF

MASTER OF SCIENCE

IN

INDUSTRIAL ENGINEERING

By

Süheyl Güleçyüz

January 2018

FINITE PERTURBATION ANALYSIS METHODS FOR OPTIMIZATION OF INVENTORY SYSTEMS WITH NON-STATIONARY MARKOV-MODULATED DEMAND AND PARTIAL INFORMATION

By Süheyl Güleçyüz

January 2018

We certify that we have read this thesis and that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Kağan Gökbayrak(Advisor)

Emre Nadar(Co-Advisor)

Seçil Savaşaneril Tüfekçi

Özlem Çavuş İyigün

Approved for the Graduate School of Engineering and Science:

Ezhan Karaşan
Director of the Graduate School

ABSTRACT

FINITE PERTURBATION ANALYSIS METHODS FOR OPTIMIZATION OF INVENTORY SYSTEMS WITH NON-STATIONARY MARKOV-MODULATED DEMAND AND PARTIAL INFORMATION

Süheyl Güleçyüz
M.S. in Industrial Engineering
Advisor: Kağan Gökbayrak
Co-Advisor: Emre Nadar
January 2018

The state of the economy may fluctuate due to several factors, and the customer demand is affected from the fluctuations of the state of the economy. Although the inventory holders can predict the state of the economy based on the demand realizations, they generally do not have the true state information. The lack of information can be extended to the transition probabilities in the state, and the demand distributions associated with each state. Further extensions may include the actual number of demand states. We consider a single-item, periodic-review inventory system with Markov-modulated discrete-valued demand, constant lead time, and full backlogging. The true demand distribution state is partially observed based on the realized demands. We study the infinite horizon average cost minimization problem, in which the optimal inventory replenishment policy is a state-dependent base-stock policy. We develop a local search method based on finite perturbation analysis (FPA) to find the base-stock levels for a finite number of discretized state beliefs. We then extend our search method to the unknown transition matrix and demand distribution case. We compare the FPA-based local search algorithm with a myopic base-stock policy, the Viterbi algorithm, and the sufficient statistics method, in terms of the average cost. Finally, we analyze how the average cost changes with respect to the estimated number of demand states when the actual number of states is unknown.

Keywords: Inventory systems, Hidden Markov models, Base-stock policy, Finite perturbation analysis, Simulation, Baum-Welch algorithm.

ÖZET

TALEP DAĞILIMININ SAKLI MARKOV MODELLERİNE BAĞLI OLARAK DEĞİŞTİĞİ DEĞİŞTİĞİ ENVANTER SİSTEMLERİ İÇİN SONLU SARSINIM ANALİZİ YÖNTEMLERİ

Süheyl Güleçyüz
Endüstri Mühendisliği, Yüksek Lisans
Tez Danışmanı: Kağan Gökbayrak
İkinci Tez Danışmanı: Emre Nadar
Ocak 2018

Ekonominin durumu birtakım etkenlere göre değişebilmekte ve müşteri talepleri ekonomideki değişimlerden etkilenmektedir. Envanter tutucuları ekonomik durumu gerçekleşen talepleri gözlemleyerek tahmin edebilmesine rağmen genellikle bu durum hakkında kesin bilgiye sahip değildir. Bilgi eksikliği, ekonomik durumdaki değişim olasılıklarını ve bu durumların her biriyle ilişkilendirilmiş talep dağılımlarını kapsayacak şekilde genişletilebilir. Daha ileri seviyelerde ise bilgi eksikliği talep dağılımı sayısını da kapsayabilir. Envanterin dönemsel olarak gözden geçirildiği, karşılanamayan taleplerin ürün tedarik edildiğinde karşılandığı, sabit tedarik süreli, dönemsel taleplerin saklı Markov modellerine bağlı olarak değiştiği, tek ürünlü bir envanter sistemini ele almaktayız. Bu sisteminde talep dağılım durumları doğrudan gözlemlenememekte, ancak gerçekleşen talep miktarlarına bağlı olarak tahmin edilebilmektedir. Uzun vadeli ortalama dönemlik maliyet problemi üzerinde çalışmakta ve bu problem için duruma bağlı taban stok seviyesinin en iyi sipariş politikası olduğunu belirtmekteyiz. Sonlu sayıda taban stok seviyesi bulmak için sarsınım analizi destekli bir yerel arama yöntemi geliştirmekteyiz. Daha sonra, geliştirdiğimiz yerel arama yöntemini geçiş matrisi ve talep dağılımının bilinmediği duruma uyarlamaktayız. Sarsınım analizi destekli yerel arama yöntemini kısa vadeli taban stok politikası, Viterbi algoritması ve yeterli istatistik yöntemi ile ortalama maliyet bakımından karşılaştırmaktayız. Son olarak, talep dağılım sayısının bilinmediği durumlarda ortalama maliyetin tahmini talep dağılım sayısına göre nasıl değiştiğini incelemekteyiz.

Anahtar sözcükler: Envanter sistemleri, Saklı Markov modelleri, Taban stok politikası, Sonlu sarsınım analizi, Benzetim, Baum-Welch algoritması.

Acknowledgement

Firstly, I would like to express my sincere gratitude to my advisor Asst. Prof. Kağan Gökbayrak and my co-advisor Asst. Prof. Emre Nadar for their endless support, guidance and patience during my graduate study. It has always been a pleasure to work with them.

I am also very grateful to Assoc. Prof. Seçil Savaşaneril Tüfekçi and Asst. Prof. Özlem Çavuş İyigün for accepting to read and review this thesis, and their invaluable comments and suggestions.

I would like to thank my fellow colleagues and officemates for all their moral support and the good times we spent together.

I also would like to thank my mother and my father, for supporting me during my whole life. Without their efforts, I would never have achieved all my accomplishments.

Finally, I would like to thank TÜBİTAK for their support during TÜBİTAK 1001 research project.

Contents

1 Introduction **1**

2 Literature Review **6**

3 Problem Formulation **15**

3.1 Known Transition and Emission Matrices 15

3.2 Unknown Transition and Emission Matrices 19

3.3 Unknown Number of Demand States 20

4 FPA-Based Local Search Method **22**

4.1 Fundamentals of FPA 22

4.2 FPA-Based Local Search Method 23

4.3 Initialization of the Base-Stock Levels 25

4.4 Fine-Tuning Base-Stock Levels via FPA 26

4.5 FPA-Based Local Search Algorithm when the Transition and Emission Matrices are Unknown 30

4.6 Summary of the FPA-Based Local Search Method 30

5 Estimating Unknown Transition and Emission Matrices 32

5.1 Forward Procedure 33

5.2 Backward Procedure 34

5.3 Determining the State Probabilities 35

5.4 Updating the Model Parameters 36

5.5 Normalization 38

5.6 Determining the Demand Sequence Probability 38

5.7 Tolerance, Maximum Number of Iterations and Initial Matrices . 39

5.8 Summary of the Baum-Welch Algorithm 39

6 Numerical Study 41

6.1 Simulation Model 42

6.2 Numerical Experiments 43

6.2.1 Known Transition and Emission Matrices 49

6.2.2 Unknown Transition and Emission Matrices 58

6.2.3 FPA in Hidden vs. Known Transition and Emission Matrices 70

6.2.4 Unknown Number of Demand States 73

7 Conclusion 75

A Tables **84**

List of Figures

3.1 The inventory position, the replenishment, and the demand realization in period t 16

4.1 Nominal and perturbed paths when the base-stock level of the demand state 1 is perturbed by $+1$ units. 27

4.2 Nominal and perturbed paths when the base-stock level of the demand state 1 is perturbed by -1 units. 28

6.1 The logical structure of the simulation model 42

6.2 Observed state beliefs for different transition matrices 46

6.3 Base-stock levels and observed state beliefs, $R = 50$ 47

6.4 Base-stock levels and observed state beliefs, $R = 200$ 47

6.5 Base-stock levels and observed state beliefs, $R = 500$ 48

6.6 95% confidence intervals for $\frac{100(\lambda_{\text{MP}}-\lambda_{\text{PA}})}{\lambda_{\text{MP}}}$ vs. R, for $N = 2$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. . 50

6.7 95% confidence intervals for $\frac{100(\lambda_{\text{MP}}-\lambda_{\text{PA}})}{\lambda_{\text{MP}}}$ vs. R, for $N = 3$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. . 51

6.8 95% confidence intervals for $\frac{100(\lambda_{\mathrm{MP}}-\lambda_{\mathrm{PA}})}{\lambda_{\mathrm{MP}}}$ vs. R, for $N = 4$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. . . 52

6.9 $\frac{100(\lambda_{\mathrm{PA}}-\lambda_{\mathrm{K}})}{\lambda_{\mathrm{K}}}$ vs. R, for $N = 2$, $R \in \{50, 200, 500\}$, , $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. 54

6.10 $\frac{100(\lambda_{\mathrm{PA}}-\lambda_{\mathrm{K}})}{\lambda_{\mathrm{K}}}$ vs. R, for $N = 3$, $R \in \{50, 200, 500\}$, , $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. 55

6.11 $\frac{100(\lambda_{\mathrm{PA}}-\lambda_{\mathrm{K}})}{\lambda_{\mathrm{K}}}$ vs. R, for $N = 4$, $R \in \{50, 200, 500\}$, , $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. 56

6.12 95% confidence intervals for $\frac{100(\lambda_{\mathrm{M}}-\lambda_{\mathrm{PA}})}{\lambda_{\mathrm{M}}}$ vs. R, $N = 3, c = 1, h = 1, b = 10$. 58

6.13 State transition probabilities vs. number of estimations, $N = 2$, $R_P \in \{7, 14\}$, max iterations 600 and 1200, respectively. 60

6.14 State transition probabilities vs. number of estimations, $N = 3$, $R_P \in \{7, 14\}$, max iterations 600 and 1200, respectively. 60

6.15 State transition probabilities vs. number of estimations, $N = 4$, $R_P \in \{7, 14\}$, max iterations 600 and 1200, respectively. 61

6.16 Demand distribution parameters vs. Number of Estimations, $N = 2$, $R_P \in \{7, 14\}$, max iterations 600 and 1200, respectively. 62

6.17 Demand distribution parameters vs. Number of Estimations, $N = 3$, $R_P \in \{7, 14\}$, max iterations 600 and 1200, respectively. 62

6.18 Demand distribution parameters vs. Number of Estimations, $N = 4$, $R_P \in \{7, 14\}$, max iterations 600 and 1200, respectively. 63

6.19 95% confidence intervals for $\frac{100(\lambda_{\mathrm{MPH}}-\lambda_{\mathrm{PAH}})}{\lambda_{\mathrm{MPH}}}$ vs. R, for $N = 2$, $R_P = 7$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. 64

6.20 95% confidence intervals for $\frac{100(\lambda_{\text{MPH}}-\lambda_{\text{PAH}})}{\lambda_{\text{MPH}}}$ vs. R, for $N = 3$, $R_P = 7$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. 65

6.21 95% confidence intervals for $\frac{100(\lambda_{\text{MPH}}-\lambda_{\text{PAH}})}{\lambda_{\text{MPH}}}$ vs. R, for $N = 4$, $R_P = 7$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. 66

6.22 95% confidence intervals for $\frac{100(\lambda_{\text{H}}-\lambda_{\text{PAH}})}{\lambda_{\text{H}}}$ vs. R, for $N = 2$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$, 30 replications, 10,000 periods 67

6.23 95% confidence intervals for $\frac{100(\lambda_{\text{H}}-\lambda_{\text{PAH}})}{\lambda_{\text{H}}}$ vs. R, for $N = 3$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$, 30 replications, 10,000 periods 68

6.24 95% confidence intervals for $\frac{100(\lambda_{\text{H}}-\lambda_{\text{PAH}})}{\lambda_{\text{H}}}$ vs. R, for $N = 4$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$, 30 replications, 10,000 periods 69

6.25 95% confidence intervals for $\frac{100(\lambda_{\text{PAH}}-\lambda_{\text{PA}})}{\lambda_{\text{PA}}}$ vs. R, for $N = 2$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. . 70

6.26 95% confidence intervals for $\frac{100(\lambda_{\text{PAH}}-\lambda_{\text{PA}})}{\lambda_{\text{PA}}}$ vs. R, for $N = 3$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. . 71

6.27 95% confidence intervals for $\frac{100(\lambda_{\text{PAH}}-\lambda_{\text{PA}})}{\lambda_{\text{PA}}}$ vs. R, for $N = 4$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. . 72

6.28 95% confidence intervals for average cost vs. the number of demand distributions, in which the number of demand distributions are unknown, for update interval 7, maximum iterations 600, $l \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. Light gray area is the confidence interval of the average cost in which the number of demand states is known and the emission and transition matrices are unknown, for $N = 4$. Dark gray area is the confidence interval of the average cost in which the true states, the emission and transition matrices and the number of demand distribution are known. 73

6.29 95% confidence intervals for $\frac{100(\lambda_{\text{UN}} - \lambda_{\text{TS}})}{\lambda_{\text{TS}}}$ vs. the number of demand distributions, in which the number of demand distributions are unknown, for update interval 7, maximum iterations 600, $l \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. Light gray area is the confidence interval of the average cost in which the number of demand states is known and the emission and transition matrices are unknown, for $N = 4$. Dark grey area is the confidence interval of the average cost in which the true states, the emission and transition matrices and the number of demand distribution is known. 74

List of Tables

3.1 Summary of notation 18

A.1 FPA-based local search algorithm calculation times for available HMM parameter information, in seconds, with the corresponding demand distributions and transition matrices for each N, $c = 1,\ h = 1,\ b = 10$. 84

A.2 95% confidence intervals for the average costs, initial base-stock levels found via the discretized myopic policy vs FPA, $N = 2$, $c = 1,\ h = 1\ b = 10$. 85

A.3 95% confidence intervals for the average costs, initial base-stock levels found via the discretized myopic policy vs FPA, $N = 3$, $c = 1,\ h = 1\ b = 10$. 86

A.4 95% confidence intervals for the average costs, initial base-stock levels found via the discretized myopic policy vs FPA, $N = 4$, $c = 1,\ h = 1\ b = 10$. 87

A.5 95% confidence intervals for the average costs, continuous myopic policy vs FPA, $N = 2$, $c = 1,\ h = 1\ b = 10$. 88

A.6 95% confidence intervals for the average costs, continuous myopic policy vs FPA, $N = 3$, $c = 1,\ h = 1\ b = 10$. 89

A.7 95% confidence intervals for the average costs, continuous myopic policy vs FPA, $N = 4$, $c = 1,\ h = 1\ b = 10$. 90

A.8 Comparison of 95% confidence intervals for the average costs obtained by Viterbi algorithm, sufficient statistic method, and FPA, $N = 3$, $c = 1,\ h = 1\ b = 10$. 91

A.9 Average cost results and 95% confidence intervals for the myopic policy, max iterations 600 and 1200 for each $R_P \in \{7, 14\}$, respectively, $N \in \{2, 3, 4\}$ $c = 1,\ h = 1\ b = 10$. 92

A.10 95% confidence intervals for the average costs, initial base-stock levels found via the disretized myopic policy vs FPA, for unknown transition matrix and demand distribution case, $N = 2$, $R_P = 7$, $c = 1,\ h = 1\ b = 10$ 93

A.11 95% confidence intervals for the average costs, initial base-stock levels found via the disretized myopic policy vs FPA, for unknown transition matrix and demand distribution case, $N = 3$, $R_P = 7$, $c = 1,\ h = 1\ b = 10$ 94

A.12 95% confidence intervals for the average costs, initial base-stock levels found via the disretized myopic policy vs FPA, for unknown transition matrix and demand distribution case, $N = 4$, $R_P = 7$, $c = 1,\ h = 1\ b = 10$ 95

A.13 95% confidence intervals for the average costs, continuous myopic policy vs FPA, for unknown transition matrix and demand distribution case, $N = 2$, $R_P = 7$, $c = 1,\ h = 1\ b = 10$ 96

A.14 95% confidence intervals for the average costs, continuous myopic policy vs FPA, for unknown transition matrix and demand distribution case, $N = 3$, $R_P = 7$, $c = 1,\ h = 1\ b = 10$ 97

A.15 95% confidence intervals for the average costs, continuous myopic policy vs FPA, for unknown transition matrix and demand distribution case, $N = 4$, $R_P = 7$, $c = 1$, $h = 1$ $b = 10$ 98

Chapter 1

Introduction

In real life, the state of the economy may fluctuate over time, and these fluctuations tend to have several effects on the factors that determine the customer demand, i.e., purchasing power, seasonality, trends, market share. Realized demand values increase or decrease in proportion to the forementioned factors. Thus, it is necessary to consider the state of the economy before making an ordering decision, to avoid or minimize the risk of the excess inventory and the unmet demands.

In most real-life inventory problems, the inventory holders do not have the information of the current economic state when they make an ordering decision. However, they can predict the current state by observing the previous and current customer demand data, i.e. previous sales. The lack of available information for the inventory holders can often be extended into the change probabilities in the economic state, and the probabilities of the customer demand values for each state. The lack of the information of the number of economic states can be considered as a further extension.

Inventories are continuous or discrete assets, i.e. raw materials, unfinished products, work-in-process materials, finished goods, which are held for sale or to be consumed in some business or production process. Holding some inventory incurs a cost per unit amount, whereas if the inventory is not held there exists a shortage cost due to possible customer dissatisfaction, delays in processes, lost sales etc. Unmet inventory can be considered as a lost sale, or backlogged after the replenishment order is placed. In order to meet the demands, items need to be ordered from one or more external suppliers and placed in inventory. Purchasing cost per amount of placed orders, cost per order, transportation cost etc. lead to an ordering cost. An order may be placed as soon as it is replenished if the lead time is zero, or at the end of a constant or stochastic lead time after the replenishment.

Ordering decision is given according to a replenishment policy which aims to determine the order amount and time that minimizes the inventory and ordering costs. The most commonly applied replenishment policies in the literature are (R, Q), (s, S), and base-stock policies. In (R, Q) policy, a fixed order amount of Q is given when the inventory position falls below the fixed re-order point R. (s, S) policy differs from (R, Q) policy such that an order amount that raises the inventory position to order-up-to-level S is replenished, when the inventory position falls below the re-order point s. Base-stock policy differs from (s, S) policy such that the replenishment order amount is specified in order to raise the inventory position to the base-stock level.

Replenishment policies can also be classified as continuous-review and periodic-review policies. The forementioned policies can be applied for both continuous-review and periodic-review cases. In continuous-review policies, the value of the inventory position is continuously reviewed, and as soon as the inventory position falls below a threshold called re-order point, the ordering decision is given. In periodic-review policies, the inventory position is reviewed once in a prespecified time interval called "period". The inventory position is reviewed in each period in order to give an ordering decision. The order is replenished in the period when the inventory position falls below the reorder point.

As previously mentioned, customer demand is subject to change according to the state of the economy. The state of the economy represents the factors that increase or decrease the customer demand, i.e. the customer need and desire for the item, seasonality, market share and competition, the economic situation (economic shortage, welfare etc.), and the purchasing power. Lower levels of the state of the economy are associated with lower demand realization values, whereas higher levels of the state of the economy are associated with higher demand realizations. Hence, different demand distributions can be associated for each state of the economy. And the demand distributions fluctuate as the state of the economy changes over time with respect to some transition probabilities. in the literature, there are several studies that model such inventory systems as a Markov decision process (MDP).

The information of the actual state of the economy and the current demand distribution is often not available, or not necessarily to be observed, for some observers, i.e. inventory holders. However, it is possible to predict the current state via the previous demand realizations. The inventory position is also fully observable. In the literature, there are studies that model these type of inventory systems as a partially observed Markov decision process (POMDP) on an uncountable state space for the demand states.

Let us give a brief information about hidden Markov models (HMM). In hidden Markov models, the actual state sequence, which is controlled by a known transition matrix, cannot be observable. However, the emission sequence, which is controlled by the emission probabilities that are the probabilities of the observations conditional on a given state, is fully observable. Then, the true state sequence can be estimated via the observable emission sequence. For the inventory systems modeled as a POMDP, the change probabilities on the demand states and the demand distributions correspond to transition and emission matrices, respectively.

In some cases, the probabilities of changes in the state of the economy and the demand distributions associated with each state are unknown. Hence, the uncertainty of information is extended into the transition matrix and the demand distributions. Unknown transition matrix and demand distributions can be estimated by using the previously observed data. Baum-Welch algorithm is the most well-known algorithm that is developed to estimate the unknown transition and emission matrices of HMM's. There are also some cases in which the actual number of different states of the economy is not available, and all calculations can be performed based on the estimated number of states.

In our study, we consider an inventory system modeled as a POMDP with several demand states. We develop an FPA-based local search method in order to minimize the average inventory and ordering cost. We also consider the different levels of available information, such that unknown transition matrix and demand distributions, and unknown number of states. Then, we observe the behavior of the FPA-based local search algorithm and compare the performances of the myopic policy and the FPA-based local search algorithm in unknown transition matrix and demand distributions case.

The FPA-based local search method calculates the differences in the average cost for any change in each base-stock level that correspond to one state, by using the previously observed data and without performing separate simulation runs for any change in each base-stock level. Then, it fine-tunes the base-stock levels by considering these differences once in every fixed update interval lengths, and finds base-stock levels that reduce the average cost. Before we perform our method, we discretize the continuous state space to deal with the uncountable number of base-stock levels.

We also compare the FPA-based local method with the myopic policy that calculates a base-stock level over the continuous state beliefs in every period, and several other methods called sufficient statistic and Viterbi algorithm.

The rest of this thesis is organized as follows: In Chapter 2, we review the literature about inventory management and optimal inventory policies, perturbation analysis, the systems which are applicable for perturbation analysis, partially observed Markov decision process (POMDP), and applications of perturbation analysis (PA) on inventory systems. In Chapter 3, we describe the inventory model that we consider, and define and formulate our problem. In Chapter 4, we briefly mention the fundamentals of FPA, then introduce our FPA-based local search method and explain the initialization and fine-tuning stages in detail. In

Chapter 5, we consider the case in which the transition matrix and the demand distribution are unknown. Then, we explain the transition matrix and demand distribution estimation procedure via Baum-Welch algorithm. In Chapter 6, we firstly introduce our simulation model. Then, we provide numerical studies for the FPA-based local search algorithm for both known and unknown transition matrix and demand distribution cases. And we compare the myopic policy and the FPA-based local search method for both cases separately, in terms of the average cost differences. We also compare the FPA-based local search method with Viterbi and Sufficient Statistics methods. Finally, we observe the relation between the estimated number of demand states and the average cost, when the number of demand states is unknown. In Chapter 7, we conclude the thesis and mention how this study can be extended as a future work.

Chapter 2

Literature Review

In this chapter, we review the literature related with the inventory systems modeled as MDP and POMDP, parameter estimation in hidden Markov models (HMM) with unknown parameters, the systems which perturbation analysis (PA) is applied on, and the applications of PA on inventory systems.

There exist several papers that consider the inventory systems with non-stationary and Markov-modulated demand. Song and Zipkin (1993) [1] consider a continuous-review inventory system with full backlogging, stochastic lead time, and Markov-modulated Poisson demand. There are linear holding and shortage costs per item, and a fixed cost per order and linear variable ordering costs per item. Their objective is to minimize the total discounted cost over both finite or infinite horizon cases. They show that the state-dependent (s, S) ordering policy is optimal when there is a positive fixed cost per order, and the state-dependent base-stock policy is optimal when the fixed ordering cost is zero.

Sethi and Cheng (1997) [2] consider a periodic-review inventory system with Markov-modulated demand including the cyclic demand model, full backlogging, and zero lead time, with convex holding and shortage costs that depend on the demand state, and fixed cost per order and linear variable ordering cost per item. They generalize the optimality of (s, S) ordering policy to finite and infinite horizon problems.

Beyer and Sethi (1997) [3] extend the study of Sethi and Cheng (1997) [2], and show that the state-dependent (s, S) policy is optimal for infinite-horizon average cost problems.

The forementioned papers consider the case in which the true demand state is observable, hence the true demand distribution is known. There exist papers in the literature that consider the inventory systems in which the current demand state cannot be observed.

Treharne and Sox (2002) [4] consider a periodic-review inventory system with constant lead time, and linear holding, shortage, variable ordering costt. They model the system as a POMDP in which the current demand state can be partially observed throughout the previous demand realizations. They show that the belief-dependent base-stock policy is optimal for the finite-horizon total cost problem.

Arifoğlu and Özekici (2010) [5] consider a periodic-review inventory system with random yield and finite capacity that operates in a random environment. The demand state depends on the environment and partially observable throughout the observations on the environment of the inventory manager. They show that the state-dependent modified inflated base-stock policy is optimal for the infinite-horizon total discounted cost problem.

Bayraktar and Ludkovski (2010) [6] consider a continuous-time inventory system modeled as an MDP. The demand state is partially observed throughout the previous demand realizations. They consider backlogging and lost sales, and derive optimal ordering policies for both cases.

We extend our literature review to parameter estimation in hidden Markov models. There are several algorithms which are developed in order to estimate unknown parameters of hidden Markov models. Viterbi (1967) [7] develops an algorithm that finds the most probable state sequence for a given observed sequence and known transition and emission matrices. The Baum-Welch algorithm is developed in order to deal with unknown emission and transition matrices and to find an estimator for unknown model parameters. It is developed for hidden Markov models as a special case of expectation maximization (EM) algorithm, which is developed to estimate the maximum likelihood for statistical models. The Baum-Welch algorithm was first described in several papers published by Baum and his colleagues in late 1960's and early 1970's; see Baum and Petrie (1966) [8], Baum and Eagon (1967) [9], Baum and Sell (1968) [10], Baum et. al. (1970) [11], and Baum (1972) [12]. Vercauteren et. al. (2007) [13] develop an online Bayesian signal processing algorithm in order to estimate the unknown transition matrix. Iki et. al. (2007) [14] develop a learning algorithm of the reward-penalty type for the communicating case of MDPs with known state and unknown transition matrix.

Ho et.al. (1979) [15] present a gradient estimation technique for the buffer storage design problem in a serial production line. They calculate the gradient vector with respect to several values of buffer sizes Ho and Cao (1983) [16] generalize the perturbation analysis on queuing networks and verify the unbiasedness of PA. Ho et. al. (1983) [17] experimentally verified the first order conditions. Cao (1985) [18], and Suri and Zazanis (1988) [19] address the consistency of PA estimators for the M/G/l queues. Ho and Cao (1985) [20], Glasserman (1988) [21], and Vakili (1989) [22] study the reformulation of IPA estimators for discontinuous performance measures. Cao (1987) [23] shows that IPA derivative with respect to the mean service time always converge as the sample path length goes to infinity. L'Ecuyer (1990) [24] compares IPA, LR (likelihood ratio), and SF (score function) methods, then show how IPA can be considered as a degenerated special case of SF and LR methods. Glasserman (1990) [25], Hu (1991) [26], and Glasserman and Yao (1991) [27] derive the conditions in which IPA can be applied to Markov chains and large queuing networks.

Finally, we review the literature dealing with the applications of perturbation analysis (PA) on inventory systems. Many papers in this literature focus on derivations of infinitesimal perturbation analysis (IPA) and smoothed perturbation analysis (SPA) estimators. Only few papers consider other approaches such as phantom-SPA, simultaneous perturbation stochastic approximation (SPSA), and finite perturbation analysis (FPA).

IPA is a sensitivity analysis technique used to estimate the effects of infinitely small changes of a control parameter on system performance measures. IPA is an on-line computation approach that finds performance estimates for nominal and perturbed paths via gradient estimators in a single simulation run, whereas several traditional techniques require separate simulation runs.

SPA estimates the effects of infinitesimal perturbations of a control parameter on the conditional expectation of a performance measure under some condition. SPA may prove useful for problems in which IPA is not sufficient for derivation of gradient estimators. One such case is when the control parameter is the replenishment decision in inventory systems.

There exists a significant body of research on (s, S) inventory systems that considers gradient estimations with respect to s and S, in order to find the pair (s, S) that minimizes the long-run average cost per period. Bashyam and Fu (1991) [28] study an application of PA to general (s, S) inventory systems in which demand arrives according to a renewal process. They consider a single-item periodic review inventory system with full backlogging and zero lead time, independent and identically distributed (i.i.d.) continuous demand, and linear holding and shortage costs. They derive PA estimators with respect to s and S, consisting of IPA and SPA components. For the same (s, S) inventory system, Bashyam and Fu (1994) [29] develop derivative estimators with respect to s and S, for both finite-horizon and infinite-horizon problems. Unlike Bashyam and Fu (1991) [28], Bashyam and Fu (1994) [29] combine PA with the conditional Monte Carlo.

Fu and Healy (1992) [30] consider an inventory model with linear ordering, holding, and shortage costs, and a fixed ordering cost. Fu and Healy (1992) [30] derive an IPA estimator with respect to s, and a PA estimator with IPA and SPA contributions with respect to $\Delta = S - s$. In addition, they compare the gradient-based perturbation analysis algorithm and the retrospective approach proposed in Healy and Schruben (1991) [31]. The retrospective approach finds s and S values that minimize the average cost per period for the realized demand values in the last n-period. They also make a comparison of these two methods in terms of their weaknesses: slow convergence for the gradient-based method and computational burdens for the retrospective approach when the horizon is too long. Fu and Healy (1997) [32] propose a hybrid algorithm based on these two methods compared in Fu and Healy (1992) [30].

Fu (1994) [33] considers an undiscounted single-item periodic review (s, S) inventory system with general i.i.d. continuous demands, full backlogging, nonnegative constant lead time, fixed setup cost, and general holding and shortage costs per item. Fu (1994) [33] derives IPA estimators with respect to s and $q = S - s$, and an additional SPA component for q, to find (s, S) pairs that minimize the long-run average cost per period.

Vázquez-Abad and Cepeda-Jüneman (1999) [34] develop the phantom-SPA method for the inventory model in Bashyam and Fu (1991) [28]. Phantom-SPA method estimates the contribution of the finite differences via parallel phantom systems which are the replicas of the nominal system, and use common random variables with the nominal system, instead of performing independent replications. They derive one-sided and two-sided SPA derivatives, in order to develop parallel phantoms. Some off-line estimations may be required to implement SPA estimators. By phantom-SPA approach, such off-line stages can be bypassed and derivatives can be estimated on-line via a single sample path.

Bashyam and Fu (1998) [35] calculate gradient estimators for (s, Q) inventory systems with random lead times and a service level constraint. They take derivatives of the long-run expected average cost per period function and the service level constraint with respect to s and Q to find IPA estimators for s and gradient estimators with IPA and SPA components for Q.

Zhang and Fu (2005) [36] derive sample path derivatives for (s, S) inventory systems with price determination. They consider a firm that makes production and pricing decisions under stationary i.i.d. demand that depends on a constant product price p, under the assumptions of zero lead time, full backlogging, variable ordering, holding, and shortage costs, and a fixed ordering cost. They derive gradient estimators with respect to control parameters s, S, and p: an IPA estimator with respect to s and PA estimators with IPA and SPA components with respect to S and p.

Karim et al. (2010) [37] consider a continuous-review (s, S) inventory system with a batch stochastic Petri net n customers and lost sales. Demand of each customer arrives in batch according to Poisson distribution. They conduct perturbation analysis for the inventory level and the stockout rate with respect to demand and replenishment rates, separately. They also conduct the sensitivity analysis for the stockout rate with respect to both demand and replenishment rates.

Several other papers study applications of PA on production-inventory systems. Bashyam et al. (1995) [38] consider a multi-product periodic review production system with no setup cost, a capacity constraint, and i.i.d. continuous demand, and switching curve policy. They assume a linear approximation for switching curve. They derive IPA estimators for a two-product system with respect to the slope of switching curve and the target base-stock levels for the two products, respectively.

Glasserman and Tayur (1995) [39] consider a single-item capacitated multi-echelon production inventory system with i.i.d. continuous demand, non-negative constant lead time, no fixed cost, linear holding and shortage costs, and a modified base-stock policy. They derive IPA derivatives with respect to the base-stock level of each stage. Then, they propose an IPA-based algorithm to find the optimal base-stock levels. They offer IPA estimations for three different cases: finite-horizon average cost, discounted total cost, and infinite-horizon average cost. Anupindi and Tayur (1997) [40] consider a single-stage continuous time multi-product production system with a service level constraint, continuous demand, and base-stock policy. The production order of each product type is determined based on a cyclic schedule. Anupindi and Tayur (1997) [40] follow the methodology in Glasserman and Tayur (1995) [39], to derive IPA estimators. Kapuscinski and Tayur (1998) [41] consider a single-item capacitated single-stage production inventory system with full backlogging, periodic demand, and linear holding and shortage costs. Similar to Glasserman and Tayur (1995) [39], the model in Kapuscinski and Tayur (1998) [41] operates under a modified base-stock policy. Kapuscinski and Tayur (1998) [41] derive IPA estimators with respect to base-stock levels across periods.

Paschalidis et al. (2001) [42] introduce a new approach that combines Large Deviations (LD) and PA techniques, in order to minimize the expected inventory costs. They apply this approach to a discrete time make-to-stock (MTS) production system that consists of one producer and one finished goods inventory with full backlogging and stationary demand. Paschalidis et al (2004) [43] propose an algorithm that combines LD and PA techniques for MTS systems with a service level constraint. They consider a single-item periodic-review supply chain production model with full backlogging, periodic demand, and production capacity. Each stage operates according to a base-stock policy. The combined approach includes IPA estimators with respect to base-stock levels across stages. Combining LD and PA leads to more accurate IPA estimations for larger range of stockout probabilities and quicker initial feasible point calculations.

Zhao and Melamed (2004) [44] apply PA to a single-stage single-item continuous review MTS production-inventory system with backorders, modeled as a stochastic fluid model (SFM) paradigm and operating under a base-stock policy. They derive IPA derivatives with respect to the base-stock level and the production rate parameter, in order to measure the sensitivity on those performance metrics. Zhao and Melamed (2004)[44] assume that the inventory system starts with the base-stock level, and that the left and right derivatives coincide. Zhao and Melamed (2006) [45] relax these two assumptions in their study. Zhao and Melamed (2007) [46] extend the analysis in Zhao and Melamed (2006) [45] to inventory sytems with lost sales. Fan et al. (2009) [47] extend the analysis in Zhao and Melamed (2006) [45] to inventory systems operating under (s, S) policy. They derive IPA derivatives with respect to order-up-to-level S and reorder point s. Melamed et al. (2010) [48] model the problem described in Zhao and Melamed (2006) [45] as a discrete MTS model, which has a notion of lead times, and maintains the identity of individual demands, replenishments, and orders, in contrast to an SFM model.

Özdemir et al. (2005) [49] develop an IPA-based solution procedure to find a policy that minimizes the expected total cost for a capacitated transshipment problem in a two-stage multi-echelon production system. The upper-echelon supplier with infinite capacity supplies multiple lower-echelon stocking locations. A stationary stochastic demand arrives in each stocking location in each period, any unsatisfied demand is backlogged, and the lead time is zero. The propose a procedure that first calculates the optimal transshipment quantities via solving a linear program developed for the problem and then updates the order-up-to-levels of each stocking location via calculating IPA gradients of the total cost function with respect to the order-up-to levels.

Chew et al. (2010) [50] consider a two-echelon periodic-review assemble-to-stock (ATS) systems full backlogging, infinite production capacity, i.i.d. demand, deterministic delivery lead time that may be different for different products, and a deterministic assembly lead time for all products. In order to find the optimal order-up-to levels for products that minimize the average total cost per period over a finite planning horizon, they develop an algorithm that combines the steepest descent algorithm and IPA estimators with respect to the order-up-to levels.

There are also papers in the literature that apply PA to supply chain models by calculating simultaneous perturbation stochastic approximation (SPSA) estimators. SPSA departs from the aforementioned techniques in that each element of the perturbation vector is generated randomly and independently of each other. Schwartz et al. (2006) [51] propose an algorithm based on SPSA, for two different supply chain models: a single-echelon supply chain operating under an Internal Model Control decision policy and a three-echelon supply chain operating under a centralized Model Prediction Control policy. The algorithm initializes the input vectors that consist of control parameters, generates a random perturbation vector by generating each element of the vector independently, evaluates the objective function for the perturbed path, and approximates the gradient. Finally, it updates the input vector by using the approximated gradient. Özgüven and Özbay (2011) [52] develop a solution procedure based on SPSA for a humanitarian inventory management and emergency logistics system.

In the literature, only few papers study the application of FPA estimators to supply chain models. FPA proceeds in a similar manner to IPA. It considers the effects of finite perturbations of control parameters on a performance measure.

Sadok et al. (2013) [53] consider a single-product manufacturing system that consists of one machine and one buffer area with a service level. Demand follows a Normal distribution and is satisfied from the buffer. Any unsatisfied demand is lost. The system is modeled as a discrete flow model. They show the unbiasedness of the estimators. They then derive FPA estimators for production, inventory, and lost sales costs, with respect to the minimum cumulative production order quantity per period, in order to minimize the total cost over a finite planning horizon.

Ayed et al. (2017) [54] derive FPA estimators for a manufacturing system that consists of one finished goods inventory and two manufacturers. One manufacturer produces a single type of product. Unsatisfied demands are lost and incurs a certain cost. Its maximum production rate is always lower than the demand rate and it is subject to random failures that increase with time and production rate. The other manufacturer is the subcontractor. It has a constant production rate, a unit production cost, and a random availability rate that follows a uniform low. In order to find the optimal production planning, they propose an algorithm that builds upon FPA estimators and Nelder-Mead algorithm, a heuristic search that is proposed by Nelder and Mead (1965) [55], is usually combined with the simulated annealing technique.

To our knowledge, there is no study in the literature that considers PA for inventory systems with demand modeled as a hidden Markov process with possibly unknown transition and/or emission matrices.

Chapter 3

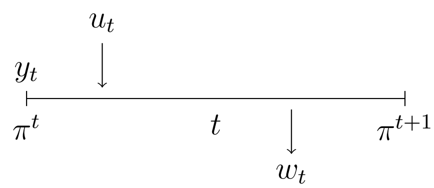
Problem Formulation

In this chapter, we define and formulate our problem, with three different levels of available information: Chapter 3.1 formulates the problem when the transition matrix, the demand distributions and the number of states are known, Chapter 3.2 formulates the problem when the transition matrix and the demand distribution are unknown and the number of demand states are known, Chapter 3.3 formulates the problem when the transition matrix, the demand distribution and the number of states are unknown.

3.1 Known Transition and Emission Matrices

We consider a single-item, discrete-time inventory system with Markov-modulated discrete-valued demand and partial information about the true demand distribution state. All unmet demand is backlogged. Replenishment lead-time L is a constant multiple of the review period. Replenishment orders u_t are placed at the beginning of each period $t + L$, $t \in \{1, 2, \dots\}$. Demand in period t arrives after the replenishment order is placed, distributed with one of N different probability distributions. We denote the demand realization in period t by $w_t \in \mathcal{M} = \{0, 1, 2, \dots, M\}$, $M \in \mathbb{Z}_+$. Figure 3.1 illustrates the inventory position, the replenishment order, and the demand realization in period t:

Figure 3.1: The inventory position, the replenishment, and the demand realization in period t



We model the demand distribution state of the world process $\{d_t\}$ as a hidden Markov chain with $N \times N$ transition matrix $P = (p_{ij})$, $i, j \in \mathcal{N} = \{1, \ldots, N\}$, where $p_{ij} = \mathbb{P}\{d_{t+1} = j | d_t = i\}$. Since $\{d_t\}$ is a hidden Markov chain, the actual state of the demand distribution in period t, $d_t \in \{1, \ldots, N\}$, is uncertain and partially observed through the demand process $\{w_t\}$. Therefore, we define the state belief vector $\pi^t = (\pi_1^t, \pi_2^t, \ldots, \pi_N^t)$ where π_i^t represents the probability that the actual state of the demand distribution is i at the beginning of period t. Note that $\pi^t \in \Pi = \{\pi \in [0,1]^N : \sum_{i \in \mathcal{N}} \pi_i = 1\}$. Hence, there exist uncountably many possible state belief vectors.

The initial state belief vector π^1 and the previous demand observations $w_1, w_2, \ldots, w_{t-1}$ form together to define the information vector. The information vector $I_t = (\pi^1, w_1, w_2, \ldots, w_{t-1})$ that contains all the available information up to period t. $\pi_i^t = \mathbb{P}\{d_t = i | I_t\}$ represents the probability that the actual state of demand distribution is i, conditional on information vector I_t, yielding a sufficient statistics.

We denote the inventory position at the beginning of period t and before the replenishment, y_t. Note that, $y_t \in \mathbb{Z}$. We also denote the replenishment order quantity in period t by $u_t \in \mathbb{Z}_+ \cup \{0\}$. We formulate our problem as a Markov decision process on the state space $\Pi \times \mathbb{Z}$ with the control process $\{u_t\}$.

Let $r_i(k) = \mathbb{P}\{w_t = k | d_t = i\}$ and $\bar{r}_\pi(k) = \mathbb{P}\{w_t = k | \pi_t = \pi\} = \sum_{i \in \mathcal{N}} \mathbb{P}\{w_t =$

$k|d_t = i\}\mathbb{P}\{d_t = i|\pi_t = \pi\} = \sum_{i\in\mathcal{N}} r_i(k)\pi_i$ denote the probability of demand realization $k \in \{0, \ldots, M\}$ when the actual demand distribution state is i, $\forall t$, and when the state belief is π, $\forall t$, respectively. We also denote the realized demands up to period t by $\omega^{t-1} = \{w_1, \ldots, w_{t-1}\}$. Then the state belief vector is updated over time as follows:

$$
\begin{aligned}
\pi_i^{t+1} &= \mathbb{P}\{d_{t+1} = i|\pi^1, \omega^{t-1}, w_t = k\} \\
&= \sum_{j\in\mathcal{N}} \mathbb{P}\{d_{t+1} = i|d_t = j, \pi^1, \omega^{t-1}, w_t = k\}\mathbb{P}\{d_t = j|\pi^1, \omega^{t-1}, w_t = k\} \\
&= \sum_{j\in\mathcal{N}} p_{ji}\frac{\mathbb{P}\{d_t = j, w_t = k|\pi^1, \omega^{t-1}\}}{\mathbb{P}\{w_t = k|\pi^1, \omega^{t-1}\}} \\
&= \frac{\sum_{j\in\mathcal{N}} p_{ji}\mathbb{P}\{w_t = k|d_t = j, \pi^1, \omega^{t-1}\}\mathbb{P}\{d_t = j|\pi^1, \omega^{t-1}\}}{\sum_{j\in\mathcal{N}} \mathbb{P}\{w_t = k|d_t = j, \pi^1, \omega^{t-1}\}\mathbb{P}\{d_t = j|\pi^1, \omega^{t-1}\}} \\
&= \frac{\sum_{j\in\mathcal{N}} p_{ji}r_j(w_t)\pi_j^t}{\sum_{j\in\mathcal{N}} r_j(w_t)\pi_j^t} \\
&= T_i(\pi^t, w_t = k), \forall t \in \mathbb{Z}_+, \forall i \in \mathcal{N}.
\end{aligned}
$$

And the inventory position is updated as follows:

$$
y_{t+1} = y_t + u_t - w_t.
$$

Table 3.1 summarizes the notation we use in our formulation.

Table 3.1: Summary of notation

Symbol	Description
N	Number of demand distributions.
$\mathcal{N}$	Finite collection of possible demand states, i.e., $\{1, 2, \dots, N\}$
M	Maximum amount of demand.
$\mathcal{M}$	Finite collection of possible demand values, i.e., $\{0, 1, \dots, M\}$
L	Lead time in terms of periods.
d_t	Demand distribution state in period t.
w_t	Demand realization in period t.
y_t	Inventory position at the beginning of period t before the replenishment.
u_t	Order quantity in period t.
Π	Uncountable collection of beliefs, i.e., $\{\pi \in [0, 1]^N : \sum_{i \in \mathcal{N}} \pi_i = 1\}$
π^t	The state belief vector in period t.
π_i	The probability that the actual state of the demand distribution in period t is i.
p_{ij}	The true transition probability from state i to state j.
P	The true transition matrix.
$r_i(.)$	The probability that the demand amount is k when the demand state is i.
$\bar{r}_\pi(.)$	The probability that the demand amount is k for a given state belief π.
E	The true emission matrix.
I_t	The information vector up to period t.
ω^t	The observed demand sequence up to period t.
c	Ordering cost per unit.
h	Holding cost per unit per period.
b	Shortage cost per unit per period.

The ordering cost in period t is linear in the order quantity and is given by cu_t. Each item held in stock incurs a unit holding cost per period h at the end of each period. Each backlogged demand incurs a unit shortage cost per period b at the end of each period. Therefore, the total inventory cost at the end of period t is defined as follows:

$$g(\pi^t, y_t{+}u_t) = \mathbb{E}\left[\max\left\{h\left(y_t + u_t - \sum_{l=0}^{L} w_{t+l}\right), b\left(-y_t - u_t + \sum_{l=0}^{L} w_{t+l}\right)\right\}\middle|\pi^t\right],$$

We assume that $c < b$. This assumption is standard in the inventory literature.

Our objective is to find an optimal inventory replenishment policy that minimizes the long-run average cost per period. We model our problem as a Markov decision process (MDP) with uncountable state space for state belief vectors, hence a partially observed Markov decision process (POMDP). Let $J^U(\pi^1, y_1)$ denote the long-run average cost per period when the initial state belief vector is π^1, the initial inventory position is y_1, and the order quantities are $U = \{u_1, u_2, \ldots\}, u_t \geq 0, t \in \{1, 2, \ldots, T\}$:

$$J^U(\pi^1, y_1) = \limsup_{T\to\infty} \frac{1}{T}\mathbb{E}\left[\sum_{t=1}^{T}[cu_t + g(\pi^t, y_t + u_t)]\middle|\pi^1, y_1\right]. \tag{3.1}$$

For our problem, the state belief-dependent base-stock policy is the optimal ordering policy [3, 4].

3.2 Unknown Transition and Emission Matrices

In this section, we extend our problem formulation for the case in which the transition and emission matrices are unknown, as well as the actual demand state. The emission matrix consists of the probability mass functions for each demand state and demand value.

Let us denote the estimated demand distribution in period t by $\hat{d}_t = \arg\max_i \mathbb{P}\{d_t = i|\hat{I}_t\}$. State transition is controlled by an estimated $N \times N$ transition matrix $\hat{P} = (\hat{p}_{ij})$, where $\hat{p}_{ij} = \mathbb{P}\{d_{t+1} = j|d_t = i, \hat{I}_t\}, i, j \in \mathcal{N} = \{1, 2, \ldots, N\}$. We denote the probability of demand realization $k \in \mathcal{M} =$

$\{0, 1, \ldots, M\}$ when the estimated demand state is i by $\hat{r}_i(k) = \mathbb{P}\{w_t = k | \hat{d}_t = i\}$, and the probability of demand realization for a given state belief π by $\widehat{r_\pi}(k) = \mathbb{P}\{w_t = k | \pi_t = \pi\} = \sum_{i \in \mathcal{N}} \mathbb{P}\{w_t = k | \hat{d}_t = i\} \mathbb{P}\{d_t = i | \hat{I}_t, \pi_t = \pi\} = \sum_{i \in \mathcal{N}} \hat{r}_i(k) \pi_i$. The emission matrix is denoted by $\hat{E} = \{\hat{r}_i(k)\}, i \in \mathcal{N}, k \in \mathcal{M}$. Then, the state belief calculation is modified as follows:

$$
\begin{aligned}
\pi_i^{t+1} &= \mathbb{P}\{d_{t+1} = i | \hat{\theta}, \omega^{t-1}, w_t = k\} \\
&= \sum_{j \in \mathcal{N}} \mathbb{P}\{d_{t+1} = i | d_t = j, \hat{\theta}, \omega^{t-1}, w_t = k\} \mathbb{P}\{d_t = j | \hat{\theta}, \omega^{t-1}, w_t = k\} \\
&= \sum_{j \in \mathcal{N}} \hat{p}_{ji} \frac{\mathbb{P}\{d_t = j, w_t = k | \hat{\theta}, \omega^{t-1}\}}{\mathbb{P}\{w_t = k | \pi^1, \omega^{t-1}\}} \\
&= \frac{\sum_{j \in \mathcal{N}} \hat{p}_{ji} \mathbb{P}\{w_t = k | d_t = j, \hat{\theta}, \omega^{t-1}\} \mathbb{P}\{d_t = j | \hat{\theta}, \omega^{t-1}\}}{\sum_{j \in \mathcal{N}} \mathbb{P}\{w_t = k | d_t = j, \hat{\theta}, \omega^{t-1}\} \mathbb{P}\{d_t = j | \hat{\theta}, \omega^{t-1}\}} \\
&= \frac{\sum_{j \in \mathcal{N}} \hat{p}_{ji} \hat{r}_j(w_t) \pi_j^t}{\sum_{j \in \mathcal{N}} \hat{r}_j(w_t) \pi_j^t} \\
&= \hat{T}_i(\pi^t, w_t = k, \hat{\theta}), \forall t \in \mathbb{Z}_+, \forall i \in \mathcal{N}.
\end{aligned}
$$

Hence, we modify equation (3.1) as follows:

$$
\hat{J}^U(\pi^1, y_1) = \limsup_{T \to \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^{T} [c u_t + g(\pi^t, y_t + u_t)] \middle| \pi^1, y_1, \hat{P}_0, \hat{E}_0 \right]. \tag{3.2}
$$

3.3 Unknown Number of Demand States

Finally, we consider a setting in which the number of demand distributions is also unknown. In this case, we replace the true number of demand distributions with the estimated number of demand distributions. We denote the estimated number of demand distributions by $\tilde{N}$. Demand distribution state transition is controlled by an estimated $\tilde{N} \times \tilde{N}$ transition matrix $\tilde{P} = (\tilde{p}_{ij})$, where $\tilde{p}_{ij} = \mathbb{P}\{d_{t+1} = j | d_t = i, \hat{I}_t\}$, $i, j \in \tilde{\mathcal{N}} = \{1, 2, \ldots, \tilde{N}\}$.

Let us denote the probability of demand realization $k \in \mathcal{M} = \{0, 1, \ldots, M\}$ when the estimated demand distribution is i by $\widetilde{r_\pi}(k) = \mathbb{P}\{w_t = k | \pi_t = \pi\} = \sum_{i \in \hat{\mathcal{N}}} \mathbb{P}\{w_t = k | \pi_t = \pi, d_t = i, \hat{I}_t\} \mathbb{P}\{d_t = i | \hat{I}_t, \pi_t = \pi\} = \sum_{i \in \hat{\mathcal{N}}} \tilde{r}_i(k)\pi_i$. Then, the state belief update equation is modified as follows:

$$\pi_i^{t+1} = \mathbb{P}\{d_{t+1}\} = \frac{\sum_{j \in \hat{\mathcal{N}}} \hat{p}_{ji} \tilde{r}_j(w_t) \pi_j^t}{\sum_{j \in \hat{\mathcal{N}}} \tilde{r}_j(w_t) \pi_j^t} = \tilde{T}_i(\pi^t, w_t = k, \hat{\theta}), \forall t \in \mathbb{Z}_+, \forall i \in \hat{\mathcal{N}}.$$

Thus, equation (3.1) is modified as

$$\tilde{J}^U(\pi^1, y_1) = \limsup_{T \to \infty} \frac{1}{T} \mathbb{E}\left[\sum_{t=1}^{T} [c u_t + g(\pi^t, y_t + u_t)] \middle| \pi^1, y_1, \hat{P}_0, \hat{E}_0 \right]. \qquad (3.3)$$

Chapter 4

FPA-Based Local Search Method

In this chapter, we first introduce the basic concepts related to finite perturbation analysis (FPA) methods, and then explain our FPA-based local search method in detail.

4.1 Fundamentals of FPA

Perturbation analysis (PA) is a performance analysis technique in discrete event dynamic systems. The key idea of PA is to calculate the effects of some change in a system parameter Θ on some performance measure $L(x(t;\Theta,\xi))$, where ξ denotes some random vector and $x(.)$ denotes the state of the system at time t that indicates a sample path that depends on Θ and ξ. We call $x(t;\Theta,\xi)$ as the nominal path. When we change Θ by $\Delta\Theta$, we obtain a perturbed sample path $x(t;\Theta+\Delta\Theta,\xi)$. Using a single sample path and a single random number stream, PA calculates the gradient for each perturbed sample path simultaneously.

The key consideration in the performance evaluation of a discrete-event system with PA is the expected value of $L(\Theta,\xi)$, which we denote by $J(\Theta)=\mathbb{E}[L(\Theta,\xi)]$. Several questions may arise in terms of unbiasedness and consistency, defined by the following expressions, respectively:

$$\mathbb{E}\left[\frac{dL(\Theta,\xi)}{d\Theta}\right] \stackrel{?}{=} \frac{d\mathbb{E}\left[L(\Theta,\xi)\right]}{d\Theta} \tag{4.1}$$

and

$$\lim_{t\to\infty}\left[\frac{dL(x(t;\Theta,\xi))}{d\Theta}\right] \stackrel{?}{=} \frac{d}{d\Theta}[\lim_{t\to\infty} L(x(t;\Theta,\xi))]. \tag{4.2}$$

Ho and Cao (1991) discuss both issues in detail, showing that unbiasedness and consistency hold for PA estimators, when the continuity holds for the performance measure L.

As mentioned in Chapter 2, IPA is the most commonly applied PA technique. Although IPA is more common in the literature, there exist several cases in which FPA has some advantages over IPA: FPA gradient $\frac{\Delta L}{\Delta\Theta}$ estimates finite differences whereas IPA gradient $\frac{dL}{d\Theta}$ estimates differentials.

FPA is an on-line computation approach: FPA provides the sensitivity information for the performance measure L with respect to changes in the control parameter Θ. Then, the control parameter is updated via gradient estimation, according to the sensitivity information provided by FPA.

The control parameter is updated as the algorithm progresses. And the algorithm continues over the updated control parameters until the optimal parameters are found.

4.2 FPA-Based Local Search Method

As discussed in Chapter 3, our performance measure is the infinite-horizon average cost per period. Recall that we mentioned that the belief-dependent base-stock policy is optimal for our problem. Hence, our control parameters are the base-stock levels for a finite number of state belief vectors. As we consider discrete-valued demand, order quantity and inventory position, we conduct a sensitivity analysis to observe the effects of one-unit finite differences in base-stock levels. We thus propose an FPA-based local search method, to calculate the optimal base-stock levels for a finite number of state beliefs.

We first introduce the finite set of state belief vectors as follows:

$$Q_n = \left\{ (q_1, q_2, \ldots, q_N) | q_i = \frac{k_i}{n}, k_i \in \mathbb{N}, \sum_{i=1}^{N} k_i = n \right\} \tag{4.3}$$

where n is the discretization level. Note that the cardinality of this set is $|Q_n|= \frac{(N+n-1)!}{(N-1)!\,n!}$ state belief vectors. State belief transitions occur from one continuous state belief vector to another continuous state belief vector. Continuous state belief vectors are assigned to the base-stock level of the nearest element, in Q_n in terms of the Euclidean distance.

The decision vector $\mathbf{s}$ consists of the base-stock levels, across all state belief vectors in Q_n. We denote the decision vector by $\mathbf{s} = \left[s_1, s_2, \ldots, s_{|Q_n|}\right]$. The neighboring vectors $\mathbf{s}_j^+$ and $\mathbf{s}_j^-$ are defined as the vectors that differ from the decision vector by $+1$ or -1, respectively, only at component $j \in \{1, \ldots, |Q_n|\}$. Hence, there exist $2|Q_n|$ neighboring vectors in total. The set of all neighboring vectors of $\mathbf{s}$ is defined by

$$\mathsf{N}(\mathbf{s}) = \bigcup_{j=1}^{|Q_n|} \left\{ \mathbf{s}_j^+, \mathbf{s}_j^- \right\} \tag{4.4}$$

Our FPA-based local search method can be divided into two main stages: Initialization of the base-stock level vector $\mathbf{s}$ (Chapter 4.3),and fine-tuning $\mathbf{s}$ via FPA technique (Chapter 4.4)

4.3 Initialization of the Base-Stock Levels

Initial base-stock levels may have a significant effect on the number of periods required for convergence. Starting with arbitrary base-stock levels may lead to a large number of periods for convergence. However, initialization of the method with base-stock levels closer to the optimal base-stock levels reduces the number of periods required for convergence, and thus the time to run the algorithm.

We choose the initial base-stock levels based on the myopic solution to our problem. The myopic base-stock levels can be obtained by focusing on a single-period problem as in the classical newsvendor setting. The base-stock level that minimizes the average cost in the newsvendor model is calculated according to the following ratio, see Stevenson (2009) [56, p. 581] and Silver et. al. (1998) [57]:

$$s^* = F^{-1}\left(\frac{b}{h+b}\right) \tag{4.5}$$

where F is the cumulative distribution function of the single-period demand if the lead time is zero. If the lead time is non-zero, F is the cumulative distribution function of the lead time demand.

We consider the discretized myopic policy. Then, we calculate the base-stock level s_j^*, $j \in \{1, \dots, |Q_n|\}$ for each discretized state belief vector $q \in Q_n$ with respect to the probability mass function $\mathbb{P}\{w_t = k\} = \bar{r}_q(k) = \sum_{i=1}^{N} r_i(k) q_i$ and the resulting cumulative distribution function. If the lead time L is non-zero, the probability mass function is calculated for the lead time demand. We then run the simulation and update the base-stock vector $\mathbf{s}$ initialized with $\mathbf{s}^* = \{s_1^*, s_2^*, \dots, s_{|Q_n|}^*\}$.

4.4 Fine-Tuning Base-Stock Levels via FPA

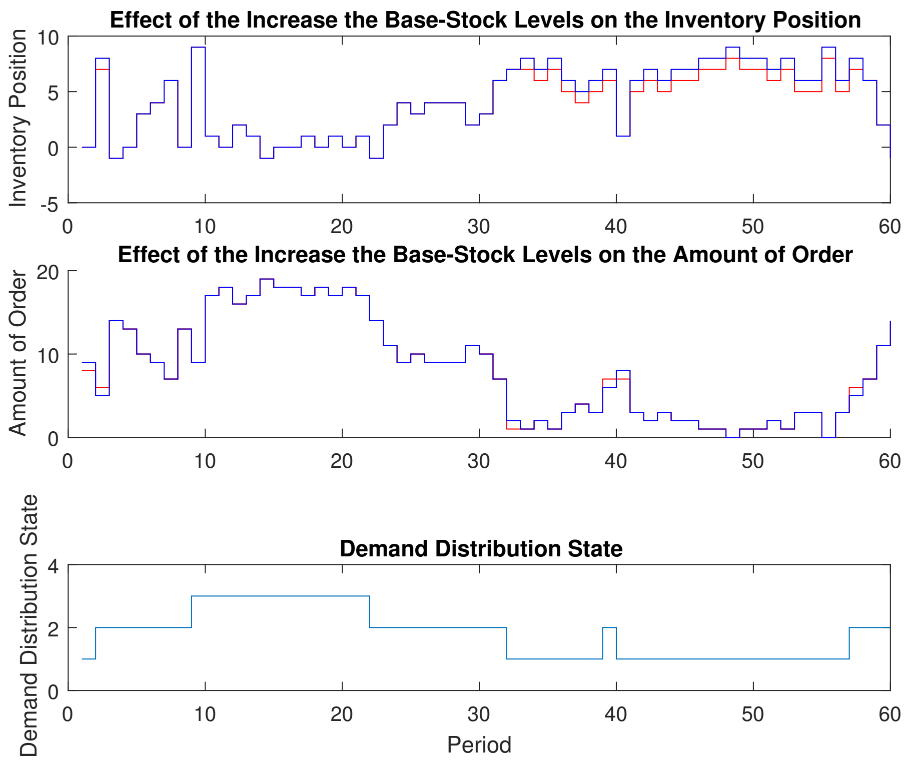
In this section we consider the case with known transition matrix and demand distributions. After determining the initial base-stock values, the next step is to improve the values of $\mathbf{s}$ with our FPA-based local search method.

FPA-based local search method calculates the differences in the average cost per period without performing separate runs for each neighboring vector $\mathbf{s_j^+}$, $\mathbf{s_j^-}$ of $\mathbf{s}$. Observed inventory levels and inventory positions, realized demands, placed orders, and state belief vectors are sufficient to perform all calculations.

The given update interval R that specifies the time interval between two consecutive updates. We choose a short update interval due to possibly limited data in reality. But, note that R must be large enough to observe a large number of different state beliefs. Thus, when R is large, any neighboring vector is more likely to lead to a non-zero change in the average cost. We calculate the difference estimations by performing a single-replication simulation run for R periods. After calculating the cost differences under the neighboring $\mathbf{s}$, we replace $\mathbf{s}$ with the neighboring vector that yields the largest difference in negative direction. The differences are set to 0 at the end of each update interval. We repeat this step K times.

Figure 4.1 shows the effect of one-unit increment on the base-stock level s_1 on the inventory position and the order amount, for an inventory system with $N = 3$, $L = 0$, $\mathbf{s} = [9, 13, 18]$, and $\mathbf{s_1^+} = [10, 13, 18]$. At the bottom, the actual state, which is not observed, is given. As shown in Figure 4.1, when the demand distribution state changes, the perturbation on the order amount no longer exists. And it starts again when the underlying demand distribution state is state 1.

Figure 4.1: Nominal and perturbed paths when the base-stock level of the demand state 1 is perturbed by +1 units.



The effect of increasing j^{th} component of $\mathbf{s}$, $j \in \{1, 2, \ldots, |Q_n|\}$, on the sample path starts when the condition $y_t \leq s_j$ holds for the first time for the current state beliefs q_j, and ends when $y_t \leq s_i, i \neq j$ holds, at the time the first order placed when the state belief vector is $q_i = q_j$.

Suppose that the effect of positive perturbation on s_j is observed in period t and $x_t = x'$, $x' \in \mathbb{Z}$. As seen in Figure 4.1, when the effect of the perturbation is observed, the perturbed path X_t^P+ holds an extra item. And it results with an increase in the inventory cost by h units when the nominal inventory level $X_t^N \geq 0$, and reduction in the shortage cost by p units when $X_t^N < 0$. Then, the difference in the inventory cost per period caused by perturbing s_j by +1 unit is given by

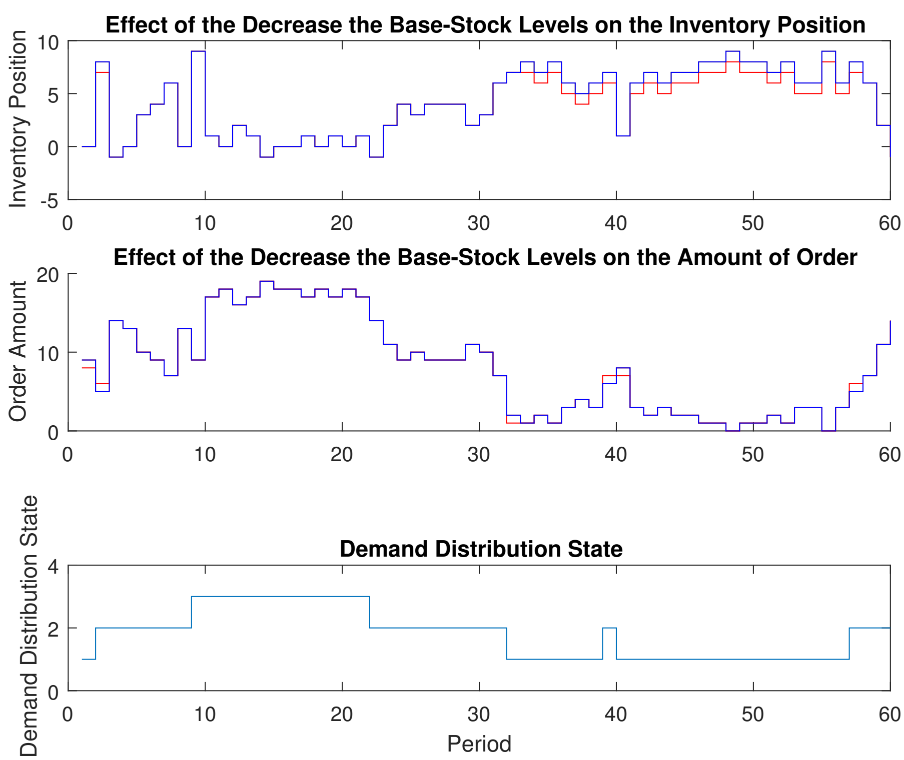
$$\Delta j^{+} = \begin{cases} h/R & x' \geq 0, \\ -b/R & x' < 0. \end{cases}$$

And the difference in ordering cost per period caused by perturbing s_j by $+1$ unit is given by

$$\Delta j^{+} = \begin{cases} c/R & \text{the first order when the state belief vector is } q_j \\ -c/R & \text{the first order after the state belief vector changes.} \end{cases}$$

Figure 4.2 illustrates the effect of one-unit decrement on the base-stock level s_1 for an inventory system with $N = 3$, $L = 0$, $\mathbf{s} = [9, 13, 18]$, and $\mathbf{s_1^-} = [8, 13, 18]$. It shows the effects on the inventory position and the order amount. At the bottom, the actual state, which is not observed, is given. Blue and red lines represent the nominal and perturbed paths, respectively. Figure 4.2 indicates that perturbing s_1 only affects the sample path in periods when $q_1 = 1$.

Figure 4.2: Nominal and perturbed paths when the base-stock level of the demand state 1 is perturbed by -1 units.



The effect of decreasing j^{th} component of $\mathbf{s}$, $j \in \{1, 2, \ldots, |Q_n|\}$, on the sample path starts with the first order when the state belief vector is q_j, and ends when the inventory position becomes less than or equal to the base-stock level that corresponds to the state belief vector q'_j such that $q'_j \neq q_j$ for the first time.

The inventory levels start to deviate from each other after different order quantities are observed. This deviation lasts until the order quantities become equal to each other. On the contrary to $X^P{+}_t$, the negative perturbed path $X^P{-}_t$ contains a missing item, hence it results with a reduction in the total inventory cost by h unit when $X_t^N \leq 1$, and increases the penalty cost by p units when $X_t^N < 0$. The difference in inventory cost per period caused by perturbing s_j by -1 unit is given by

$$\Delta j^- = \begin{cases} -h/R & x' > 0, \\ b/R & x' \leq 0. \end{cases}$$

And the difference in ordering cost per period caused by perturbing s_j by -1 unit is given by

$$\Delta j^- = \begin{cases} -c/R & \text{the first order when the state belief vector is } q_j \\ c/R & \text{the first order after the state belief vector changes.} \end{cases}$$

As seen from the equations above and Figures 4.1 and 4.2, the differences in ordering costs cancel each other when the system moves from one state belief vector to another. We thus focus on the difference in inventory costs.

In cases where the effect of perturbation is seen, consistency holds for the FPA estimator, which is given by

$$\left(\frac{\Delta J}{\Delta s_j}\right)_{FPA} = \sum_{t:X_t>0} \frac{h(X_t + \Delta s_j) - h(X_t)}{\Delta s_j} - \sum_{t:X_t<0} \frac{p(-X_t - \Delta s_j) - p(-X_t)}{\Delta s_j} \quad (4.6)$$

$\forall j \in Q_n$, and where X_t is the inventory level at the end of period t.

4.5 FPA-Based Local Search Algorithm when the Transition and Emission Matrices are Unknown

When the true transition matrix and demand distributions are unknown, we need to modify FPA-based local search method so as to estimate and update both unknown parameters alongside the base-stock vector. In addition to updating the base-stock levels in every R periods, we update the estimated transition matrix once in every R_P periods, and the estimated demand distributions once in every R_E periods, by using Baum-Welch algorithm. Note that R, R_P and R_E need not be equal to each other.

If the number of demand distributions N is also unknown, we consider an estimated number of demand distributions $\tilde{N}$ and iterate our algorithm with respect to $\tilde{N}$.

4.6 Summary of the FPA-Based Local Search Method

Our FPA-based local search method can be summarized as follows:

Step 1. Discretize the continuous state space with respect to the discretization level n.

Step 2. Determine the initial base-stock levels.

Step 3. Determine a sufficiently short update interval R and an update number K that is sufficient for convergence.

Step 4. Run the simulation model for R periods. Calculate the gradients for all neighboring vectors simultaneously.

Step 5. If there is a neighboring vector $\mathbf{s}' \in \mathsf{N}(\mathbf{s})$ with less cost, replace $\mathbf{s}$ with $\mathbf{s}'$ that yields the least cost. Otherwise, keep $\mathbf{s}$.

Step 6. If $t < RK$, go back to Step 4. Otherwise, terminate the algorithm.

Chapter 5

Estimating Unknown Transition and Emission Matrices

In this chapter, we consider the case in which the transition and emission matrices are unknown. Then, we describe the methods that we use to estimate and update the transition and emission matrices, which correspond to the state transition probabilities and the demand distributions, respectively. Recall that the emission matrix consists of the probability mass functions for each demand state and demand value, as mentioned in Chapter 3.

Let us introduce the initial hidden Markov model parameters $\hat{\theta}_0 = (\hat{P}_0, \hat{E}_0, \pi^1)$ in which the initial transition and emission matrices are denoted by $\hat{P}_0$ and $\hat{E}_0$, respectively. Starting with the initial model parameters, we estimate the model parameters $\hat{\theta} = (\hat{P}, \hat{E}, \pi^1)$ by observing the demand realizations up to period t, ω^t. We define an information vector $\hat{I}_t = (\hat{\theta}_0, \omega^t)$, yielding a sufficient statistic in order to estimate the model parameters $\hat{\theta}$.

In order to estimate the transition and emission matrices, we first initialize the emission and transition matrices arbitrarily. We update the emission matrix once in every $R_E \geq 1$ periods, and the transition matrix once in every $R_P \geq 1$ periods, respectively, by using the Baum-Welch algorithm and taking into account the realized demand values from the first period to the current period, with respect to the current state belief. Note that R_P and R_E may or may not be equal.

We next explain the Baum-Welch algorithm and its implementation in our problem. Let $\hat{\theta} = (\hat{P}, \hat{E}, \pi)$ denote the hidden Markov model with $\pi = \pi^t$, and $\hat{P} = \{\hat{p}_{ij}\}$ and $\hat{E} = \{\hat{r}_i(w_t)\}$, which correspond to the transition and emission matrices in period t, respectively. Recall that $\hat{p}_{ij} = \mathbb{P}\{d_{t+1} = j | d_t = i, \hat{I}_t\}$ denotes the transition probability from the demand states i to j, and $\hat{r}_i(w_t)$ denotes the conditional probability of demand realization w_t when the estimated demand state is i. As mentioned in Chapter 3, the demand sequence in the first t periods is denoted by $\omega^t = \{w_1, w_2, \ldots, w_t\}$. The Baum-Welch algorithm aims to find an estimator $\hat{\theta}^*$ that maximizes the probability of the demand sequence ω^T, given the model $\hat{\theta}$, that is $\hat{\theta}^* = \arg\max_{\hat{\theta}} \mathbb{P}\{\omega^T | \hat{\theta}\}$.

A single iteration of the Baum-Welch algorithm consists of several steps: Forward and backward procedures, state probability transition updates, re-estimation of emission matrix, and normalization.

Before forward and backward procedures, we initialize the hidden Markov model parameters $\hat{\theta}_0 = (\hat{P}_0, \hat{E}_0, \pi^1)$ arbitrarily. Then we set $\hat{\theta} = \hat{\theta}_0$. For the forward and backward procedures, we assume that demand values ω^T are given.

5.1 Forward Procedure

Forward procedure recursively calculates the joint probability of the realized demand sequence up to period t that ends in state i, given the hidden Markov model $\hat{\theta}$. This probability is called the forward variable:

$$\alpha(i, t) = \mathbb{P}\{\omega^t, d_t = i | \hat{\theta}\}. \tag{5.1}$$

We initialize the forward variable in period 1 as

$$\alpha(i,1) = \pi_i^1 \hat{r}_i(w_1),\ \forall i \in \mathcal{N}, \tag{5.2}$$

then we iterate $\alpha(i,t+1)$ by performing forward recursion:

$$\begin{aligned}
\alpha(i,t+1) &= \mathbb{P}\{\omega^t, w_{t+1}, d_{t+1} = i\} \\
&= \mathbb{P}\{w_{t+1} | d_{t+1} = i, \omega^t\} \mathbb{P}\{d_{t+1} = i, \omega^t\} \\
&= \hat{r}_i(w_{t+1}) \sum_{j=1}^{N} \mathbb{P}\{d_{t+1} = i | d_t = j, \omega^t\} \mathbb{P}\{\omega^t, d_t = j\} \\
&= \hat{r}_i(w_{t+1}) \sum_{j=1}^{N} \hat{p}_{ji} \alpha(j,t), \qquad (5.3) \\
&\forall i \in \mathcal{N},\ t \in \{2, 3, \ldots, T\}.
\end{aligned}$$

Finally, the probability of the demand sequence given hidden Markov model parameters $\hat{\theta}$ is given by, in terms of the value of the forward variable in period T

$$\mathbb{P}\{\omega^T | \hat{\theta}\} = \sum_{i=1}^{N} \alpha(i,T). \tag{5.4}$$

5.2 Backward Procedure

Backward procedure recursively calculates the probability of the remaining demand sequence that starts from state i, from period $t+1$ until the end of the horizon, T, given state i, and the hidden Markov model $\hat{\theta}$:

$$\beta(i,t) = \mathbb{P}\{w_{t+1}, w_{t+2}, \ldots, w_T | d_t = i, \hat{\theta}\}. \tag{5.5}$$

We initialize the backward variable in period 1 is as

$$\beta(i, T) = 1, \forall i \in \mathcal{N}, \tag{5.6}$$

then we iterate $\beta(i, t)$ by performing backward recursion:

$$\begin{aligned}
\beta(i, t) &= \mathbb{P}\{w_{t+1}, w_{t+2}, \ldots, w_T | d_t = i\} \\
&= \sum_{j=1}^{N} \mathbb{P}\{w_{t+1}, w_{t+2}, \ldots, w_T, d_{t+1} = j | d_t = i\} \\
&= \sum_{j=1}^{N} \mathbb{P}\{w_{t+2}, \ldots, w_T | d_{t+1} = j\} \mathbb{P}\{w_{t+1} | d_{t+1} = j\} \mathbb{P}\{d_{t+1} = j | d_t = i\} \\
&= \sum_{j=1}^{N} \beta(j, t+1) \hat{r}_j(w_{t+1}) \hat{p}_{ij}, \qquad (5.7) \\
&\forall i \in \mathcal{N},\ t \in \{T-1, T-2, \ldots, 1\}.
\end{aligned}$$

Finally, we find the probability of the demand sequence given $\hat{\theta}$ is calculated in terms of the backward variable as

$$\mathbb{P}\{\omega^T | \hat{\theta}\} = \sum_{i=1}^{N} \beta(i, 1) \pi_i^1 \hat{r}_i(w_1). \tag{5.8}$$

5.3 Determining the State Probabilities

The next step is to determine the state probabilities. For a given hidden Markov model $\hat{\theta}$ and demand sequence ω^T, the probability of being in state $i, i \in \mathcal{N}$, is defined as follows:

$$\gamma(i,t) = \mathbb{P}\{d_t = i|\omega^T, \hat{\theta}\}. \tag{5.9}$$

We can rewrite equation (5.9) in terms of the forward and backward variables, which are defined by equation (5.1) and (5.5), as follows:

$$\begin{aligned}
\gamma(i,t) &= \frac{\mathbb{P}\{d_t = i, \omega^T|\hat{\theta}\}}{\mathbb{P}\{\omega^T|\hat{\theta}\}} \\
&= \frac{\mathbb{P}\{\omega^t, d_t = i|\}\mathbb{P}\{w_{t+1}, \ldots, w_T|d_t = i,\}}{\mathbb{P}\{\omega^T\}} \\
&= \frac{\alpha(i,t)\beta(i,t)}{\mathbb{P}\{\omega^T\}} \\
&= \frac{\alpha(i,t)\beta(i,t)}{\sum_{i=1}^{N} \alpha(i,t)\beta(i,t)} \qquad (5.10)
\end{aligned}$$

5.4 Updating the Model Parameters

After calculating the state probabilities, the algorithm estimates the hidden Markov model parameters $\bar{P}$ and $\bar{E}$. For a given hidden Markov model $\hat{\theta}$ and demand sequence ω^T, the joint probability of being in state i in period t and state j in period $t+1$ is defined by

$$\xi(i,j,t) = \mathbb{P}\{d_t = i, d_{t+1} = j|\omega^T, \hat{\theta}\}. \tag{5.11}$$

Equation (5.11) can be rewritten in terms of the forward and backward variables as follows:

$$
\begin{aligned}
\xi(i,j,t) &= \frac{\mathbb{P}\{d_t = i, d_{t+1} = j, \omega^T | \hat{\theta}\}}{\mathbb{P}\{\omega^T | \hat{\theta}\}} \\
&= \frac{\mathbb{P}\{w_{t+2}, \ldots, w_T | d_{t+1} = j, d_t = i, w_{t+1}, \omega^T\}\mathbb{P}\{w_{t+1} | d_{t+1} = j\}}{\mathbb{P}\{\omega^T\}} \\
&\qquad \cdot \frac{\mathbb{P}\{d_{t+1} = j | d_t = i, \omega^T\}\mathbb{P}\{d_t = i, \omega^T\}\mathbb{P}\{\omega^T\}}{\mathbb{P}\{\omega^T\}} \\
&= \frac{\alpha(i,t)\beta(j,t+1)\hat{p}_{ij}\hat{r}_j(w_{t+1})}{\mathbb{P}\{\omega^T\}} \\
&= \frac{\alpha(i,t)\beta(j,t+1)\hat{p}_{ij}\hat{r}_j(w_{t+1})}{\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha(i,t)\beta(j,t+1)\hat{p}_{ij}\hat{r}_j(w_{t+1})}.
\end{aligned} \tag{5.12}
$$

Note that, $\gamma(i,t) = \sum_{j=1}^{N} \xi(i,j,t), \forall t \in 1, \ldots, T, i \in \mathcal{N}$.

The expected number of transitions from state i is given by

$$
\sum_{t=1}^{T} \gamma(i,t), \tag{5.13}
$$

and the expected number of transitions from state i to j is given by

$$
\sum_{t=1}^{T-1} \xi(i,j,t). \tag{5.14}
$$

transition probabilities from state i to j are updated as

$$
\bar{p}_{ij} = \frac{\sum_{t=1}^{T-1} \xi(i,j,t)}{\sum_{t=1}^{T-1} \gamma(i,t)}, \tag{5.15}
$$

where $i, j \in \mathcal{N}$. The emission matrix entries are also updated as

$$
\bar{r_j}(k) = \frac{\sum_{t=1, w_t = k}^{T-1} \gamma(j,t)}{\sum_{t=1}^{T-1} \gamma(j,t)}, \tag{5.16}
$$

where $j \in \mathcal{N}$, $k \in \mathcal{M}$.

Note that the initial demand state probabilities are given by

$$\bar{\pi}_i = \gamma(i, 1). \tag{5.17}$$

5.5 Normalization

Note that the forward and backward variables α and β can approach to zero as t gets larger. In order to avoid this problem, we need to normalize α and β. First, we define the normalization factor that we use for both α and β:

$$Z(t) = \sum_{i=1}^{N} \alpha(i, t). \tag{5.18}$$

We then define the normalized forward and backward variables, respectively: $\alpha^*(i, t) = \alpha(i, t)/Z(t)$, and $\beta^*(i, t) = \beta(i, t)/Z(t)$. Finally, we replace α and β with α^* and β^* for the initialization and recursion procedures in equations (5.3) and (5.4), and (5.7) and (5.8), respectively. Then, we normalize equations (5.4) and (5.8) as

$$-\log\left[\mathbb{P}\{\omega^T|\theta\}\right] = -\sum_{t=1}^{T} \log[Z(t)]. \tag{5.19}$$

5.6 Determining the Demand Sequence Probability

Recall that equation (5.4) gives the probability of the demand sequence, given $\hat{\theta}$, and equation (5.18) and (5.4) becomes equal when all demand realizations are observed. Equation (5.19) calculates the log-likelihood of equation (5.18).

Our aim to estimate the hidden Markov model parameters that maximizes the probability of the given demand sequence ω^T given $\hat{\theta}$. We try to maximize the log-normal of the demand sequence probability, until there is no further improvement on equation (5.19) between two consecutive iterations, or at least one of the conditions described in Chapter (5.7) are met.

5.7 Tolerance, Maximum Number of Iterations and Initial Matrices

There are two important parameters called "tolerance" and "maximum number of iterations" that determine the accuracy of the estimation of the model parameters. We run the algorithm, until either the difference between the probabilities of demand sequences in consecutive iterations drop below the desired tolerance level or we reach the prespecified maximum number of iterations. Lower tolerance levels often lead to more accurate model parameter estimations. However, lower tolerance levels may also lead to longer computation times.

5.8 Summary of the Baum-Welch Algorithm

The working principles of the Baum-Welch algorithm can be summarized as follows:

Step 1. Calculate $\alpha(i,t)$ and $\beta(i,t)$ via the forward and backward procedure, respectively.

Step 2. Calculate the demand state probabilities $\gamma(i,t)$ and $\xi(i,t)$.

Step 3. Update state transition and emission matrices.

Step 4. Determine the demand sequence probability.

Step 5. If the difference between the demand sequence probabilities calculated in the last two iterations are above the prespecified tolerance level or the current number of iterations is below the prespecified maximum number of iterations, go to Step 1. Otherwise, terminate the algorithm.

Chapter 6

Numerical Study

In this chapter, we conduct several numerical experiments to evaluate the performance of our FPA-based local search method for finding an optimal ordering policy in a discretized state space, in order to minimize the infinite-horizon average cost per period. For this purpose, we construct a simulation model in MATLAB for the inventory problem in Chapter 3.

We observe the performance of our method with respect to three different values for each parameter: number of demand distributions N, grid size n, update interval R, and lead time L. For the case in which the transition matrix and demand distribution information is unavailable, we have two additional sets of parameters: the transition matrix and demand distribution estimation interval R_P, and the initial transition matrix and demand distributions R_E. We also observe the effect of level of information by comparing the results of both known and unknown transition matrix and demand distribution cases. Finally, we compare our method with various methods that do not aim to find base-stock levels on a discretized state space: the myopic policy, the sufficient statistics method, and Viterbi algorithm.

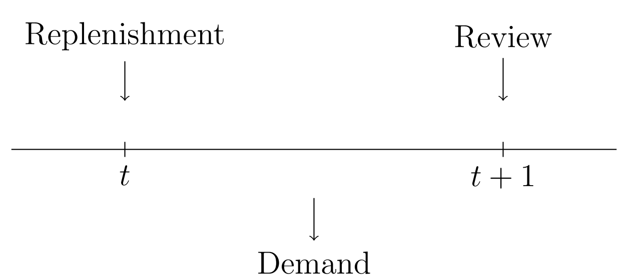
6.1 Simulation Model

We begin with the logical structure of our simulation model. The model consists of five main events: order arrival, demand realization, inventory review, and initialization and termination events.

Initialization event occurs at the beginning of the simulation run, which corresponds to the beginning of period 1. Similarly, termination event occurs at the end of the simulation run, which corresponds to the end of period T. We choose T so as to obtain a sufficiently narrow confidence interval that is enough to yield accurate simulation results. We have a short warm-up period that guarantees the elimination of the inconsistent data and the noises at the beginning of the simulation run.

Order arrival event occurs at the beginning of period t if an order is given in period $t - L$. Demand realization event occurs randomly, after the order arrival event in period t, with respect to one of N demand distributions. We assume that demands arrive in batch and once in each period. Order in period t is given before the demand is observed if the inventory position at the beginning of period t, y_t, is less than the corresponding base-stock level in period t. Inventory review event occurs at the end of every period. The cost in period t is calculated based on the inventory level at the time of the inventory review event. Figure 6.1 summarizes the logical structure of our simulation model:

Figure 6.1: The logical structure of the simulation model



Our simulation model is given as follows:

$$\begin{aligned}
\min\ & J^U(\pi^1, y_1) \\
& w\ F_i \\
x_{t+1} &= x_t + u_{t-L} - w_t \\
y_t &= x_{t-1} + \sum_{i=1}^{L-1} u_{t-i}
\end{aligned}$$

where F_i represents the cumulative distribution function of the demand distribution associated with state i.

We implement our local search method and construct a simulation model using MATLAB. The local search method calculates the initial base-stock levels via the myopic policy applied on a discretized state space, finds the base-stock levels using FPA via a single-replication simulation run and updates. The simulation model performs simulation runs for average cost calculation for the base-stock levels found for each discretized state belief in the first part with various number of replications.

6.2 Numerical Experiments

Example System

We consider instances with 2, 3, and 4 distribution states. For $N = 2$, the transition matrix is given by

$$P_2 = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix},$$

for $N = 3$, the transition matrix is given by

$$P_3 = \begin{bmatrix} 0.9 & 0.1 & 0 \\ 0.05 & 0.9 & 0.05 \\ 0 & 0.1 & 0.9 \end{bmatrix},$$

and for $N = 4$, the transition matrix is given by

$$P_4 = \begin{bmatrix} 0.9 & 0.1 & 0 & 0 \\ 0.05 & 0.9 & 0.05 & 0 \\ 0 & 0.05 & 0.9 & 0.05 \\ 0 & 0 & 0.1 & 0.9 \end{bmatrix}.$$

For $N = 2$, the two demand distributions that we consider are Binomial with parameters 20 and 0.1, and Binomial with parameters 20 and 0.9. For $N = 3$, the three demand distributions that we consider are Binomial with parameters 20 and 0.1, Binomial with parameters 20 and 0.5, and Binomial with parameters 20 and 0.9. And for $N = 4$, the four demand distributions that we consider are Binomial with parameters 20 and 0.1, Binomial with parameters 20 and 0.35, Binomial with parameters 20 and 0.65, and Binomial with parameters 20 and 0.9.

We take the ordering cost as 1, the holding cost as 1, and the shortage cost as 10.

In our study, we assume that a period is a day.

FPA-Based Local Search Method

For the FPA-based local search method, we perform simulation runs for various base-stock level update intervals. The update interval length should be sufficiently large to allow for improvements on base-stock levels. The update interval length should also be sufficiently small for rapid improvements. We consider the following base-stock level update interval lengths in our study: 50, 200, and 500 periods. Then, by taking the base-stock levels obtained via these update intervals at the end of 10,000 periods as the input, we perform 30 replications simulation runs to calculate a confidence interval for the average cost.

We work with grid sizes 4, 8, and 16. Note that if the grid size is large, update interval R should be large enough to observe a sufficient number of occurrences of different state belief vectors in our FPA-based local search method.

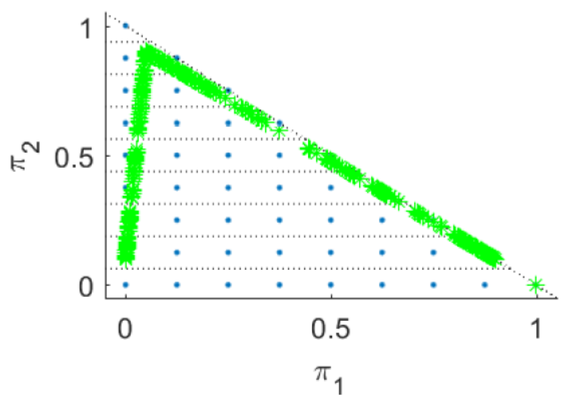
We also perform simulation runs for lead times of 0, 1, and 2 periods. We calculate the lead time demand distribution as the sum of $L+1$ Binomial distributions, according to the following discrete distribution convolution formula:

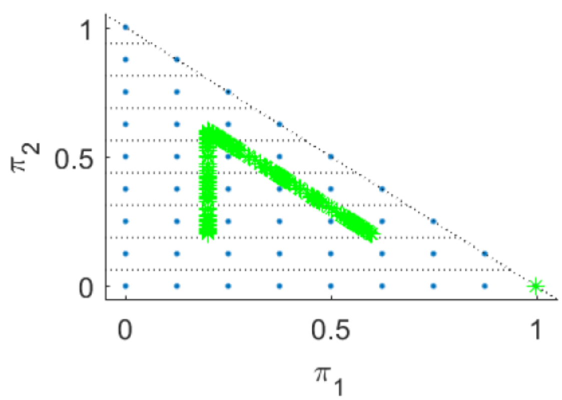
$$\mathbb{P}\{Z = z\} = \sum_{k=0}^{z} \mathbb{P}\{X = k\}\mathbb{P}\{Y = z - k\} \tag{6.1}$$

where $z \in \{0, \ldots, (L+1)M\}$, X represents single-period demand, and Y represents L-period demand and calculated iteratively.

Recall that the state-belief update is dependent on the transition matrix, observed state beliefs are affected by using different transition matrices. Figure 6.2 shows the 2 different examples for $N = 3$ and $n = 8$:

Figure 6.2: Observed state beliefs for different transition matrices





We consider 2 different transition matrices to obtain the plots in Figure 6.2:

$$A = \begin{bmatrix} 0.9 & 0.1 & 0 \\ 0.05 & 0.9 & 0.05 \\ 0 & 0.1 & 0.9 \end{bmatrix}, \text{ and } B = \begin{bmatrix} 0.6 & 0.2 & 0.2 \\ 0.2 & 0.6 & 0.2 \\ 0.2 & 0.2 & 0.6 \end{bmatrix}.$$

Figure 6.2 shows that the observed state beliefs depend on the transition matrix and most of the state beliefs are not observed. Hence, we do not need to consider the base-stock levels of the state-beliefs that are not observed.

We now present the initial base-stock levels and the base-stock levels obtained via our method in Figure 6.3, 6.4, and 6.5:

Figure 6.3: Base-stock levels and observed state beliefs, $R = 50$

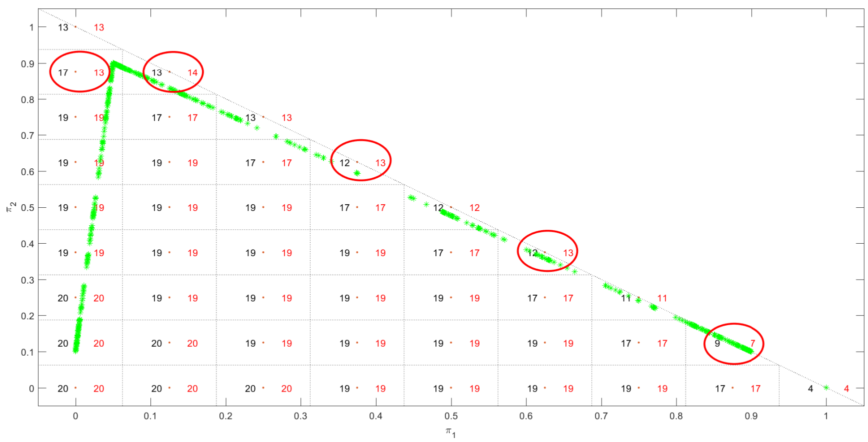


Figure 6.4: Base-stock levels and observed state beliefs, $R = 200$

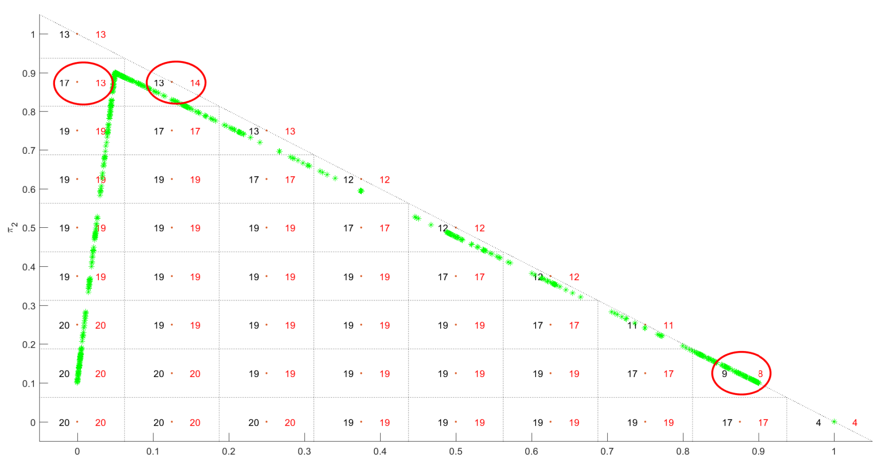
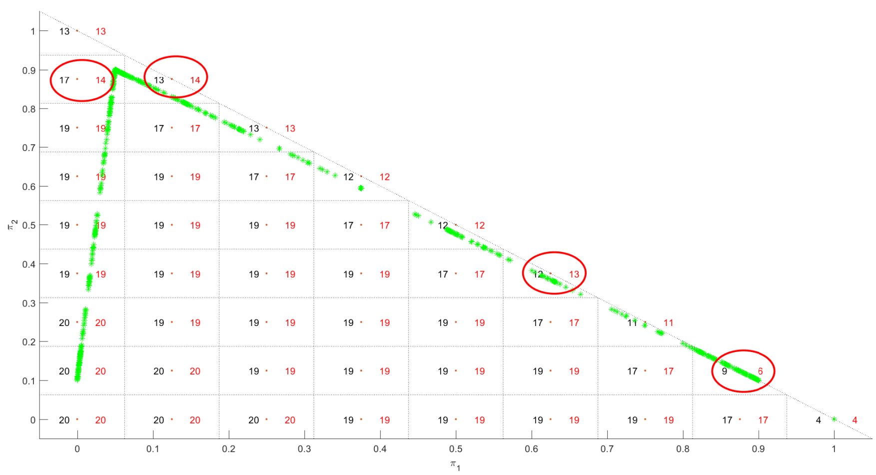


Figure 6.5: Base-stock levels and observed state beliefs, $R = 500$



The FPA-based local search method only changes the base-stock levels of the observed discretized state-beliefs. Since the FPA-based local search algorithm updates the base-stock levels using the observed data and state-belief information, the base-stock levels corresponding to a-other discretized state-beliefs remain unchanged. Note that the base-stock levels marked as black represent the initial levels, and the base-stock levels marked as red are the base-stock levels obtained by the FPA-based local search algorithm. The areas marked with red circles indicate the state beliefs and their corresponding base-stock levels, in which the base-stock level change is occurred by the FPA-based local search method. Different update intervals may lead to changes in different base-stock levels. Hence, differences in the average cost may occur between different update intervals, according to the changed base-stock levels. Note that the base-stock levels in Figure 6.3, 6.4, and 6.5 are obtained via our method are calculated over a single replication. And the green dots represent all continous state-beliefs observed in all 30 replications, in total, there are 300,000 dots in Figure 6.3, 6.4, and 6.5.

6.2.1 Known Transition and Emission Matrices

FPA vs Discretized Myopic Policy

We present the comparison of the 95% confidence intervals of the average costs of the simulation runs using the initial base-stock levels obtained via the discretized myopic policy and the simulation runs using the base-stocks obtained via the FPA-based local search method, which are denoted by λ_{MP} and λ_{PA}, respectively. We calculate the confidence intervals for the average cost for each $R \in \{50, 200, 500\}$, with respect to each different combination of the number of demand states $N \in \mathcal{N}$, $n \in \{4, 8, 16\}$, and $L \in \{0, 1, 2\}$.

Table A.2, A.3, and A.4 in Appendix show the 95% confidence intervals for the average cost when $N = 2$, $N = 3$, and $N = 4$, respectively.

Figure 6.6 illustrates the 95% confidence intervals of the improvement percentages obtained by 30 replications with 10,000 periods when we fine-tune the initial base-stock levels via the FPA-based local search method, for $N = 2$:

Figure 6.6: 95% confidence intervals for $\frac{100(\lambda_{\text{MP}}-\lambda_{\text{PA}})}{\lambda_{\text{MP}}}$ vs. R, for $N = 2$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$.

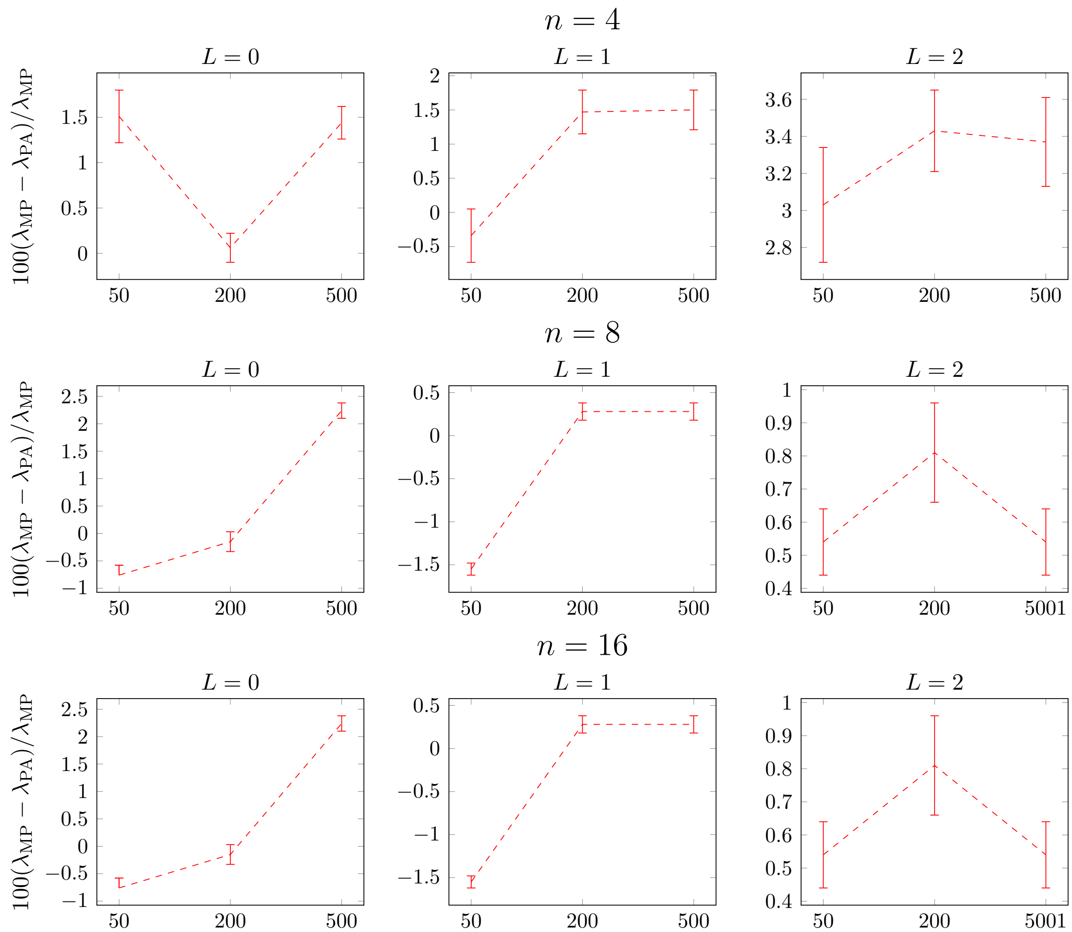


Figure 6.7 illustrates shows the 95% confidence intervals of the improvement percentages obtained by 30 replications with 10,000 periods when we fine-tune the initial base-stock levels via the FPA-based local search method, for $N = 3$:

Figure 6.7: 95% confidence intervals for $\frac{100(\lambda_{\text{MP}}-\lambda_{\text{PA}})}{\lambda_{\text{MP}}}$ vs. R, for $N = 3$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$.

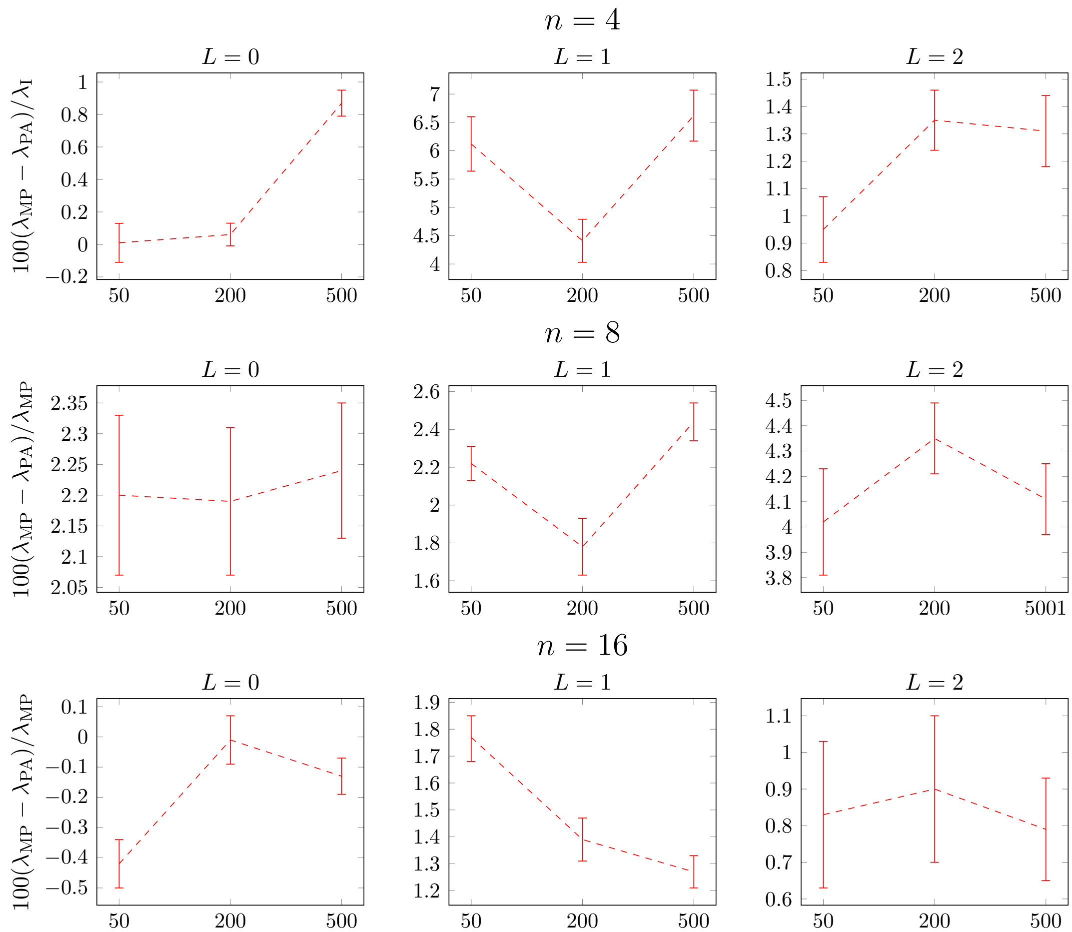
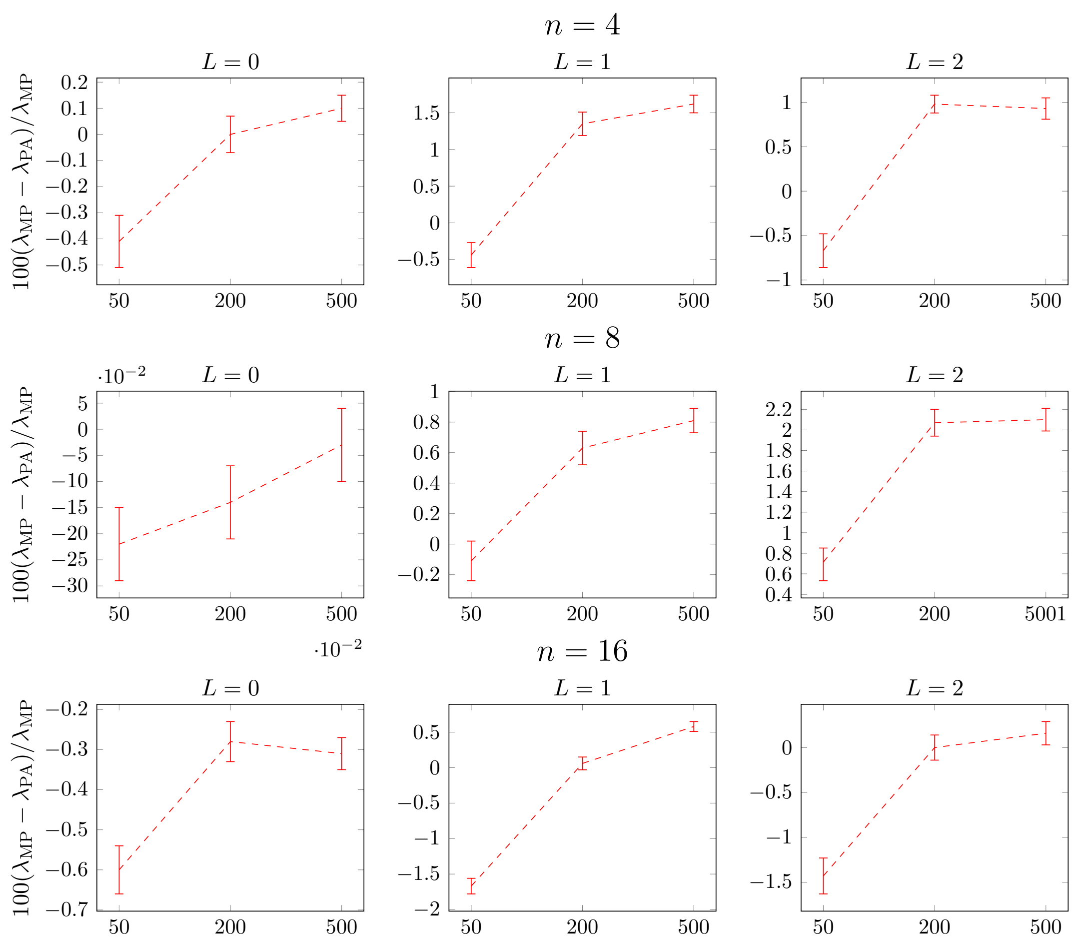


Figure 6.8 illustrates the 95% confidence intervals of the improvement percentages obtained by 30 replications with 10,000 periods when we fine-tune the initial base-stock levels via the FPA-based local search method, for $N = 4$:

Figure 6.8: 95% confidence intervals for $\frac{100(\lambda_{\text{MP}}-\lambda_{\text{PA}})}{\lambda_{\text{MP}}}$ vs. R, for $N=4$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h=1$, $c=1$, $b=10$.



Since we start with very good initial base-stock levels, improvement percentages are low, in the best case not more than 7%, and in some cases fine-tuning with the FPA-based local search method fail to reduce the average cost, as the case in which $N=4$, $L=0$, and $n=16$. Hence, the FPA-based local search method cannot perform better than the discretized myopic policy in all cases, and yields slightly less average cost in the cases it performs better.

FPA vs Continuous Myopic Policy

We now consider the continuous myopic policy. In the continuous myopic policy, we find an optimal base-stock level using the lead time demand of the current continuous state belief in every period, according to the newsvendor formula given by equation (4.5). We use the same number of demand distributions, transition matrices, demand distributions, and cost parameters that we consider in our experiments for the FPA-based local search method. We calculate the average cost by performing a simulation run that consists of 30 replications with 10,000 periods.

Figure 6.9, 6.10, and 6.11 illustrate the 95% confidence intervals of the percentages of the differences in the average cost results between the continuous myopic policy and FPA-based local search method. Note that λ_{K} and λ_{PA} denote the average cost obtained by the myopic policy and the FPA-based local search algorithm in the case in which the transition matrices and demand distributions are known.

Figure 6.9: $\frac{100(\lambda_{\text{PA}}-\lambda_{\text{K}})}{\lambda_{\text{K}}}$ vs. R, for $N = 2$, $R \in \{50, 200, 500\}$, , $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$.

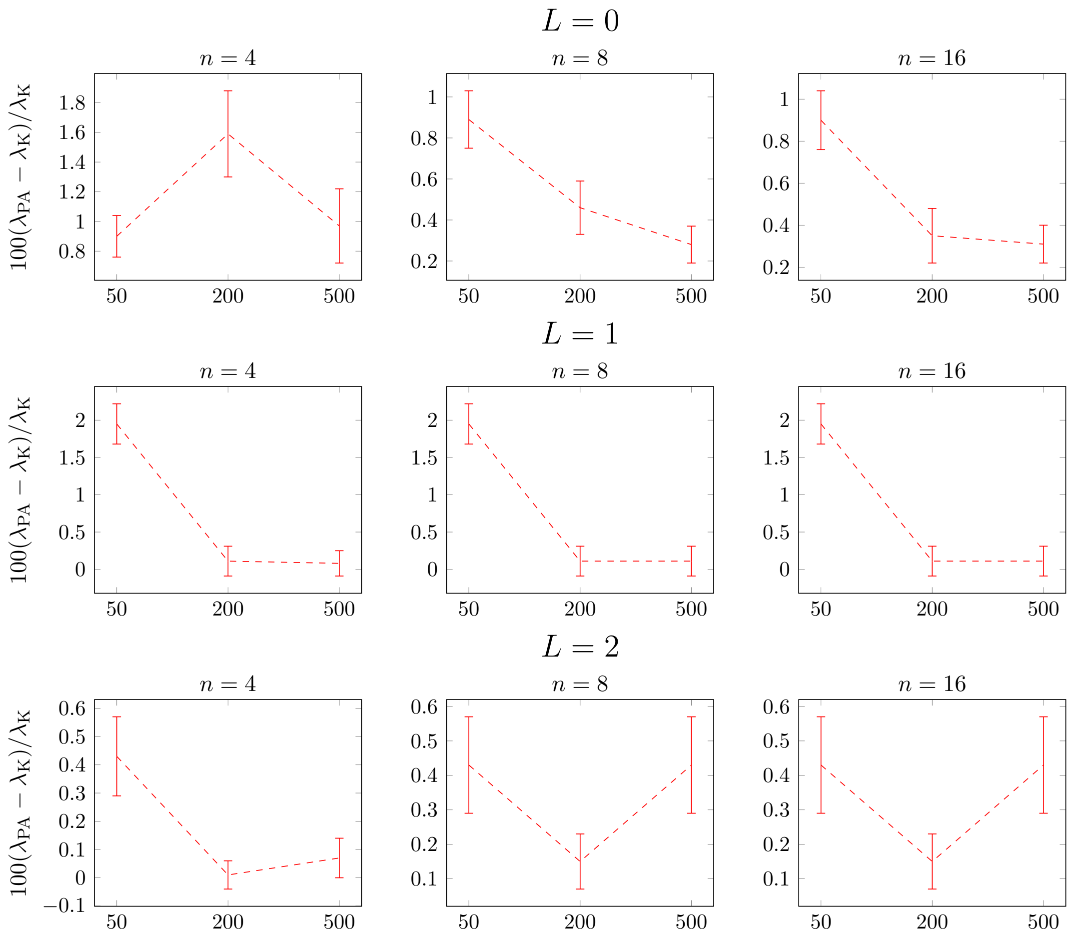


Figure 6.10: $\frac{100(\lambda_{\mathrm{PA}}-\lambda_{\mathrm{K}})}{\lambda_{\mathrm{K}}}$ vs. R, for $N = 3$, $R \in \{50, 200, 500\}$, , $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$.

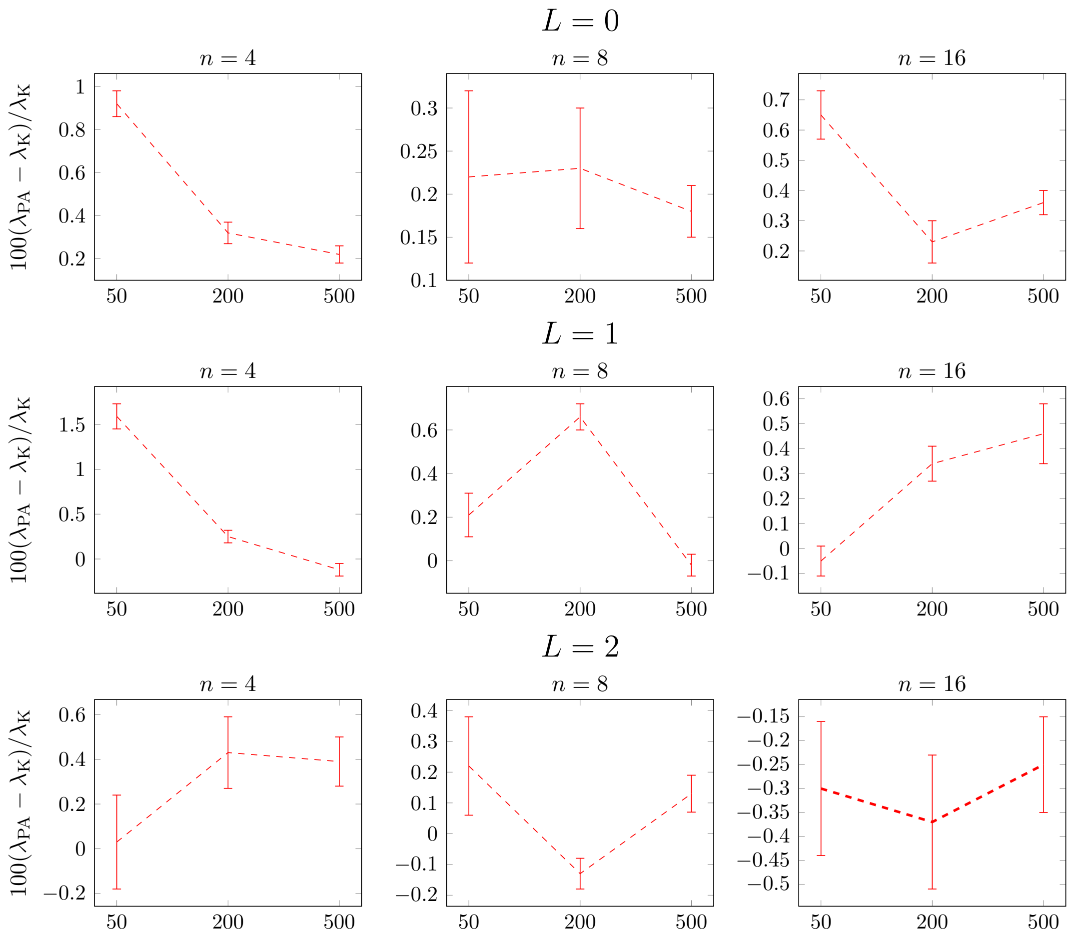
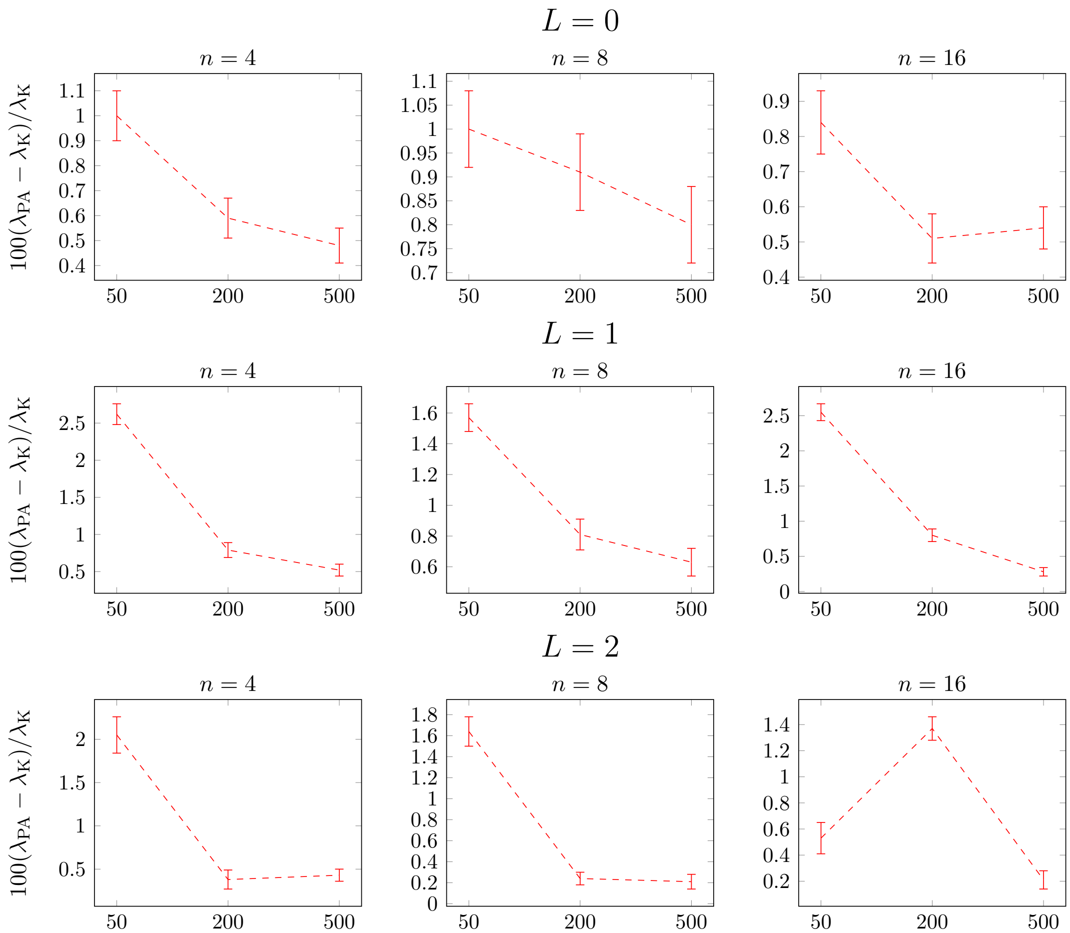


Figure 6.11: $\frac{100(\lambda_{\text{PA}}-\lambda_{\text{K}})}{\lambda_{\text{K}}}$ vs. R, for $N=4$, $R \in \{50, 200, 500\}$, , $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h=1$, $c=1$, $b=10$.



As seen in Figures 6.9, 6.10, and 6.11, although the FPA-based local search method approaches to the continuous myopic policy in terms of the average cost, even less than 1% percent difference in most cases, we cannot conclude that it performs better than the continuous myopic policy. Among the 81 different combination of N, n, L, and R, there are only 5 cases such that the FPA-based local search method perfoms better than the myopic policy: the cases when $N = 3, L = 1, n = 4, R = 200$, and when $N = 3, L = 2, n = 8, R = 200$, and all three different update intervals when $N = 3, L = 2, n = 16$. In most cases, the continuous myopic policy performs slightly better.

FPA vs Viterbi Algorithm

Viterbi (1967) propose a recursive dynamic programming algorithm that is commonly used in hidden Markov models. The algorithm estimates the most likely demand state sequence up to period t in a recursive way, by using the given known transition matrix P, conditional probability distributions of each demand realization, $r_j(k), \forall j \in \mathcal{N}$, and a given demand realization sequence up to period t, ω^t.

Let us denote the probability that the demand state in period 1 is $i \in \mathcal{N}$, and the demand realization in the first period is $k \in \mathcal{M}$ by $\nu_i^1 = r_i(k)\pi_i^1$. Suppose that ν_j^{t-1} is calculated for each demand state j, $j \in \mathcal{N}$. The probability of the most likely demand state sequence that produces the past demand realization sequence and ends in the demand state i in period t with the demand realization k, is given by

$$\nu_i^t = r_i(k)\max_j\{p_{ji}\nu_j^{t-1}\}.$$

Then, the most likely demand state in period t is defined as follows:

$$\bar{d}_t = \arg\max_i\{\nu_i^t\}.$$

FPA vs Sufficient Statistic Method

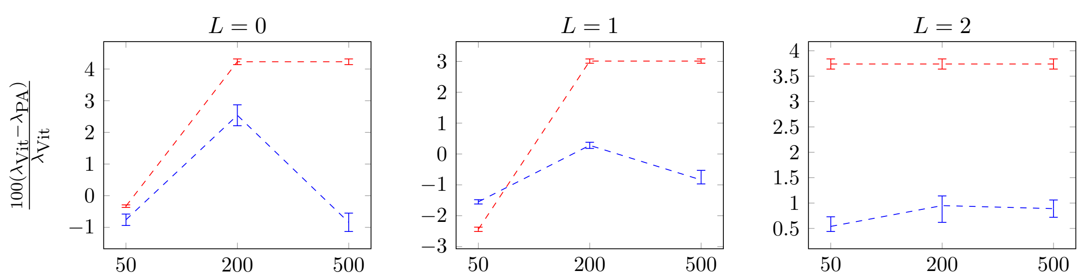
Recall that we mentioned in Chapter 3 that the state belief π yields a sufficient statistic. In the sufficient statistic method, the demand state in period t is estimated with respect to the demand probabilities. The demand state with the largest probability is assumed the most likely demand state, and assigned as the estimated demand state as follows:

$$\bar{d}_t = \arg\,\max_i\{\pi_i^t\}.$$

Table A.8 gives the comparison of the average cost values obtained by the FPA-based local search method, and Viterbi and the sufficient statistic methods, respectively:

Figure 6.12 illustrates the 95% confidence intervals of the differences in the average cost results between the FPA-based local search method, and Viterbi (red) and the sufficient statistic (blue) methods, which are provided in Table A.8. Note that λ_M corresponds to the average cost values obtained by both methods:

Figure 6.12: 95% confidence intervals for $\frac{100(\lambda_{\mathrm{M}}-\lambda_{\mathrm{PA}})}{\lambda_{\mathrm{M}}}$ vs. R, $N = 3, c = 1, h = 1, b = 10$.



The FPA-based local search method performs better improvements on Viterbi algorithm compared to the sufficient statistic method, up to 4 percent. The sufficient statistic method can be improved by at most about 2 percent, and in several cases our method performs worse than the sufficient statistic method.

6.2.2 Unknown Transition and Emission Matrices

When the information of the emission and transition matrices is not available, we also perform the transition matrix and demand distribution estimation, during the average cost calculation.

For $N = 2$, $N = 3$, and $N = 4$, we start with the $N \times N$ initial guess transition matrices $\hat{P}_N = (\hat{p}_{ij})$. For $N = 2$, the initial guess transition matrix is given by

$$\hat{P}_2 = \begin{bmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{bmatrix},$$

for $N = 3$, the initial guess transition matrix is given by

$$\hat{P}_3 = \begin{bmatrix} 1/4 & 3/8 & 3/8 \\ 3/8 & 1/4 & 3/8 \\ 3/8 & 3/8 & 1/4 \end{bmatrix},$$

and for $N = 4$, the transition matrix is given by

$$\hat{P}_4 = \begin{bmatrix} 1/5 & 4/15 & 4/15 & 4/15 \\ 4/15 & 1/5 & 4/15 & 4/15 \\ 4/15 & 4/15 & 1/5 & 4/15 \\ 4/15 & 4/15 & 4/15 & 1/5 \end{bmatrix},$$

and the demand distributions for $N = 2$, $N = 3$ and $N = 4$ are Binomial with parameters (M, p_i), where $M = 20$ and $p_i = (i - 0.5)/N$, $\forall i \in \{1, \ldots, N\}$. We choose $p_{ii} < 1/n$ to increase the difference between the initial guess transition probabilities and the actual transition probabilities. We do not choose uniform transition matrices as the initial guess matrices in order to avoid the instant convergence of initial guess transition matrix, hence in order to avoid uniform state beliefs.

In our study, we consider the following estimation interval lengths: 7 and 14 periods, i.e., one or two weeks. We assume $R_P = R_E$. We take the number of maximum iterations in a single estimation

Figure 6.13, 6.14, and 6.15 show the convergence of the state transition probabilities as the number of estimations increases, for $N = 2$, $N = 3$, and $N = 4$.

Blue lines represent the case in which $R_P = 7$, and red lines represent the case in which $R_P = 14$.

Figure 6.13: State transition probabilities vs. number of estimations, $N = 2$, $R_P \in \{7, 14\}$, max iterations 600 and 1200, respectively.

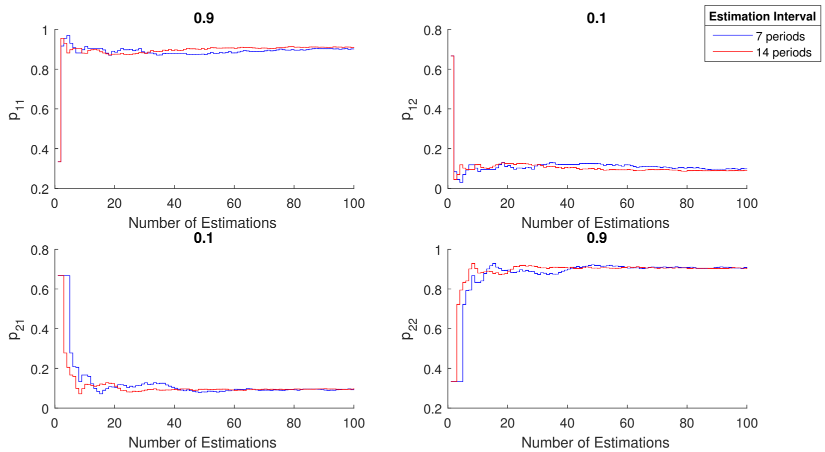


Figure 6.14: State transition probabilities vs. number of estimations, $N = 3$, $R_P \in \{7, 14\}$, max iterations 600 and 1200, respectively.

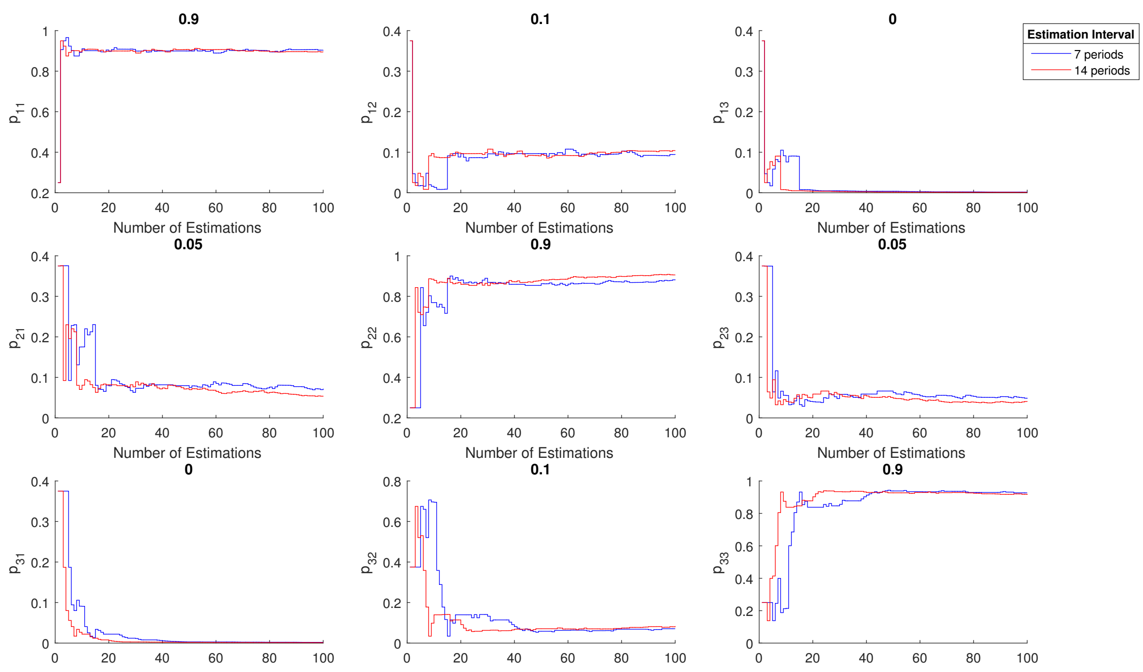


Figure 6.15: State transition probabilities vs. number of estimations, $N = 4$, $R_P \in \{7, 14\}$, max iterations 600 and 1200, respectively.

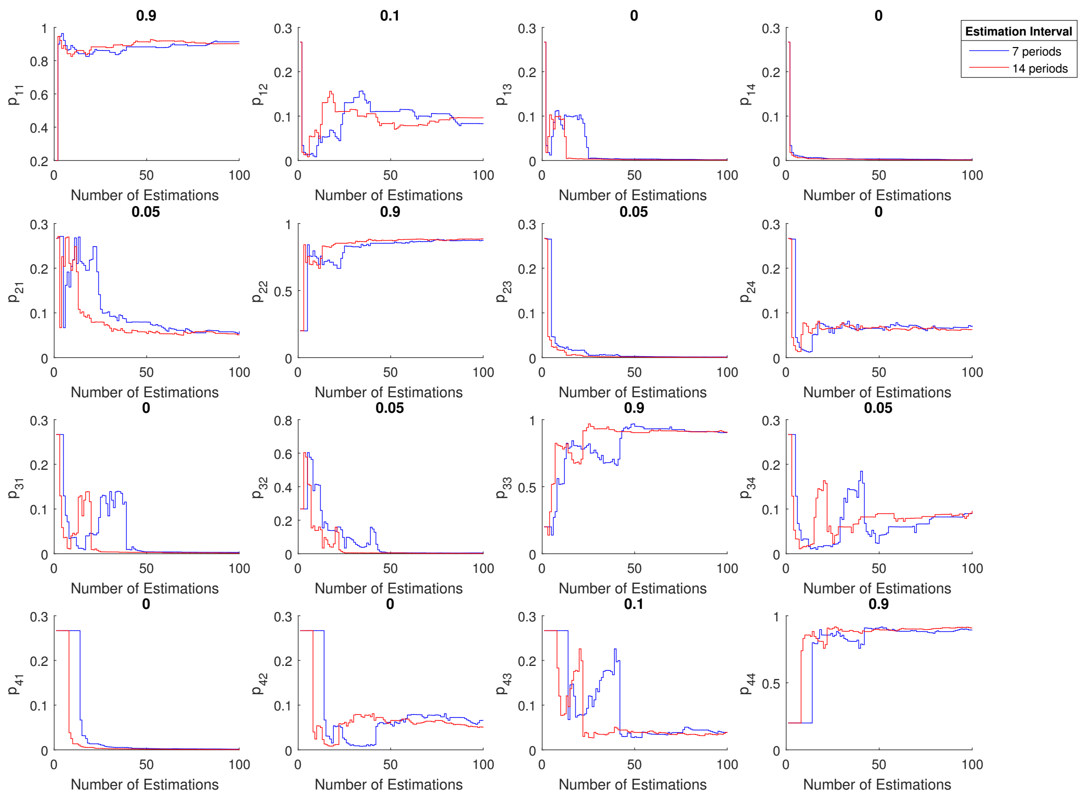


Figure 6.16, 6.17, and 6.18 ilustrate the demand distribution estimations for $N = 2$, $N = 3$, and $N = 4$, respectively:

Figure 6.16: Demand distribution parameters vs. Number of Estimations, $N = 2$, $R_P \in \{7, 14\}$, max iterations 600 and 1200, respectively.

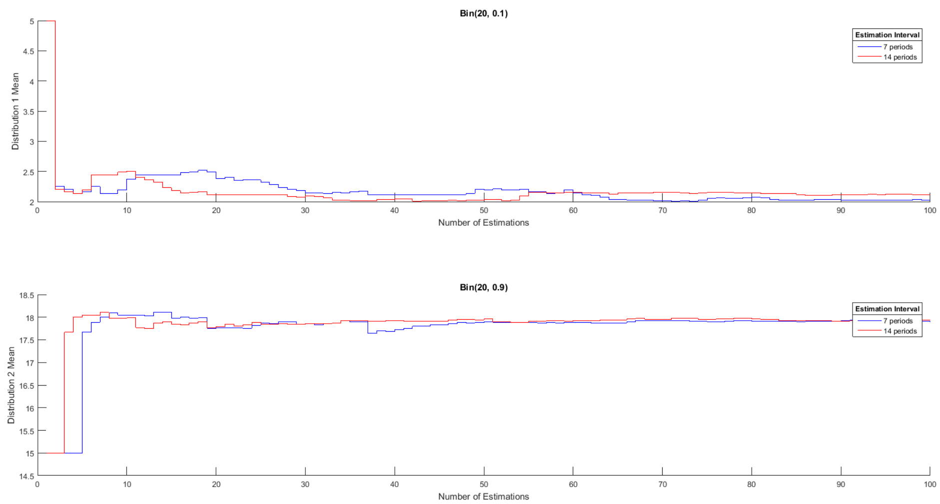


Figure 6.17: Demand distribution parameters vs. Number of Estimations, $N = 3$, $R_P \in \{7, 14\}$, max iterations 600 and 1200, respectively.

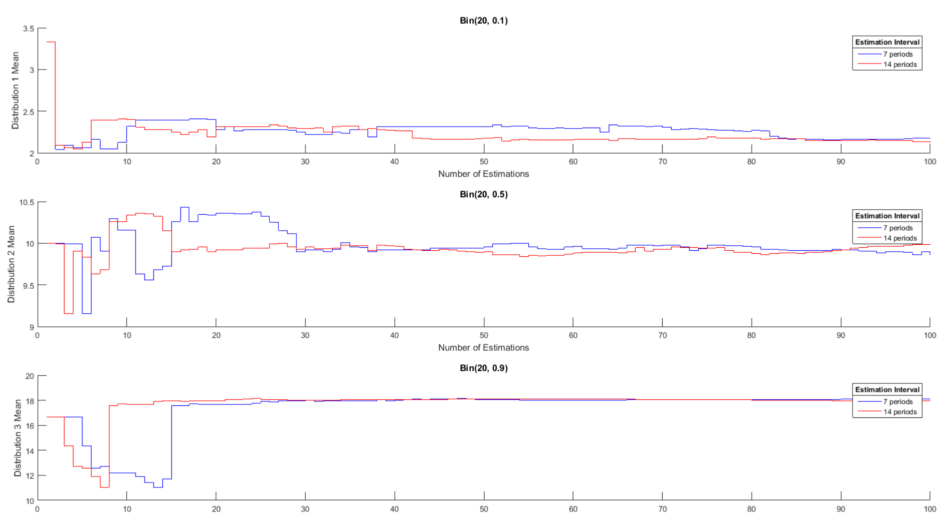
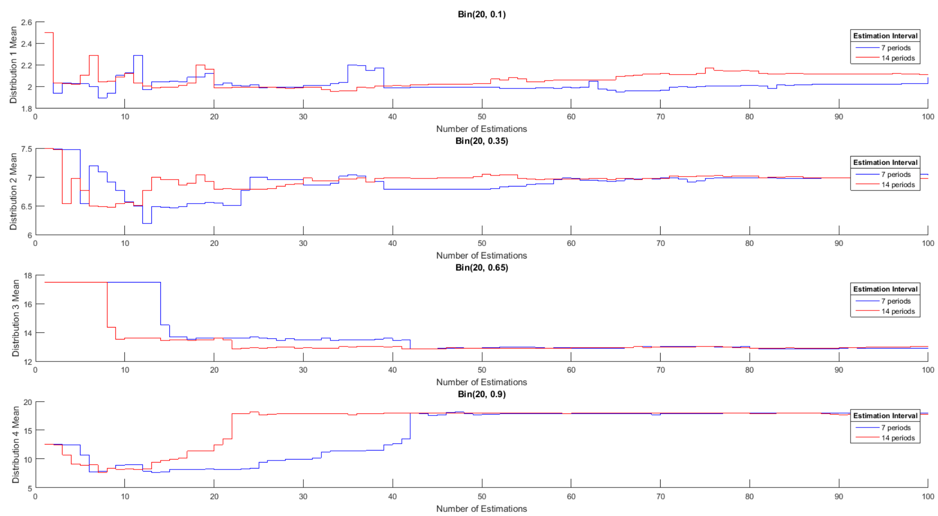


Figure 6.18: Demand distribution parameters vs. Number of Estimations, $N = 4$, $R_P \in \{7, 14\}$, max iterations 600 and 1200, respectively.



It can be seen that smaller estimation intervals require larger number of estimations to converge in probability and demand distribution parameters. If estimation interval gets smaller, the Baum-Welch algorithm trains and estimates the matrices with a very small number of observed data, at the beginning of the simulation run.

FPA vs. Discretized Myopic Policy

We now examine the performance of the FPA-based local search algorithm when the transition matrix and the demand distributions are unknown. We consider the same parameter as the previous experiments. We apply the FPA-based local search algorithm to the initial base-stock levels obtained by using the initial guess matrices.

For the unknown transition and emission matrices case, the average cost of the simulation runs using the initial base-stock levels obtained via the discretized myopic policy, and the average cost of the simulation runs using the base-stock levels obtained via the FPA-based local search method are denoted by λ_{MPH} and λ_{PAH}, respectively.

Figures 6.19, 6.20, and 6.21 illustrate the 95% confidence intervals of the differences between the simulation runs by using the initial base-stock levels The columns labeled as "Initial" correspond to the average costs obtained by the simulation runs using the initial base-stock levels, whereas in other columns, R represent the update intervals:

Figure 6.19: 95% confidence intervals for $\frac{100(\lambda_{\text{MPH}}-\lambda_{\text{PAH}})}{\lambda_{\text{MPH}}}$ vs. R, for $N=2$, $R_P=7$, $R\in\{50,200,500\}$, $n\in\{4,8,16\}$, $L\in\{0,1,2\}$, $h=1$, $c=1$, $b=10$.

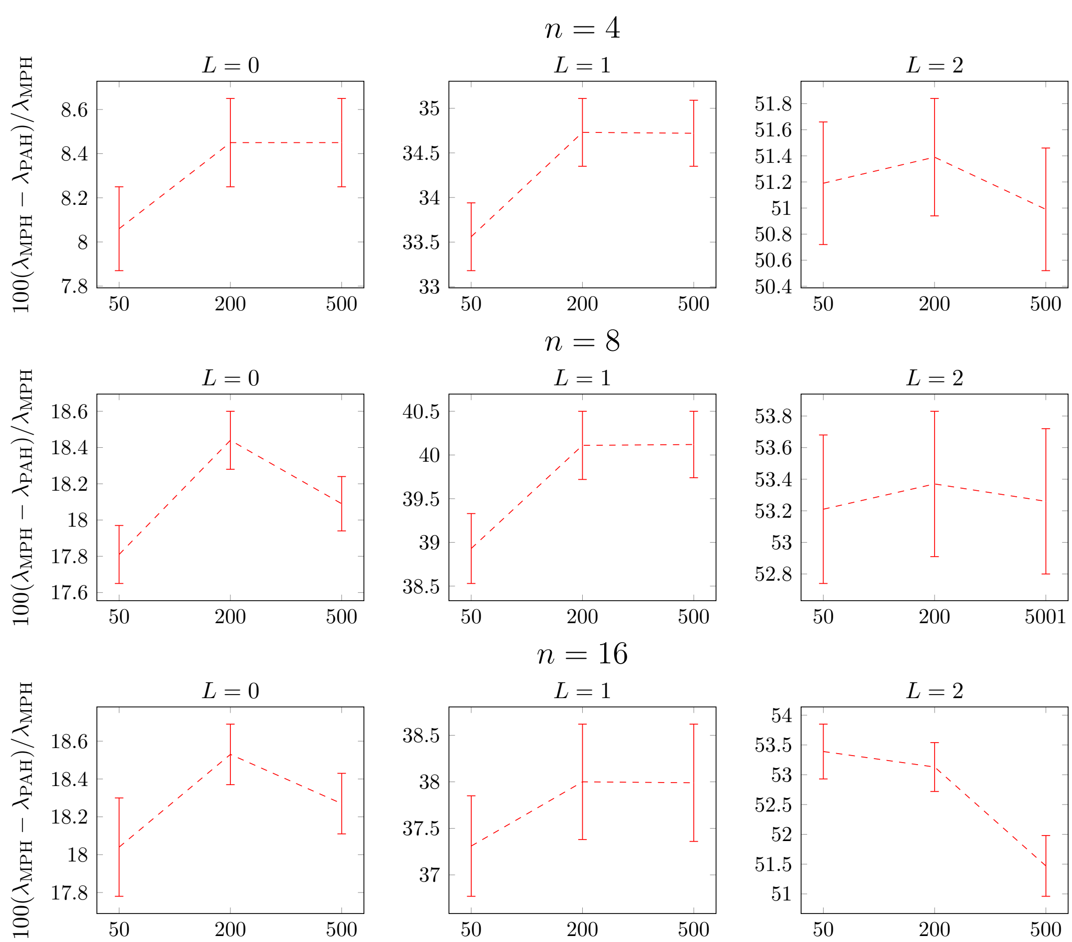


Figure 6.20: 95% confidence intervals for $\frac{100(\lambda_{\mathrm{MPH}}-\lambda_{\mathrm{PAH}})}{\lambda_{\mathrm{MPH}}}$ vs. R, for $N = 3$, $R_P = 7$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$.

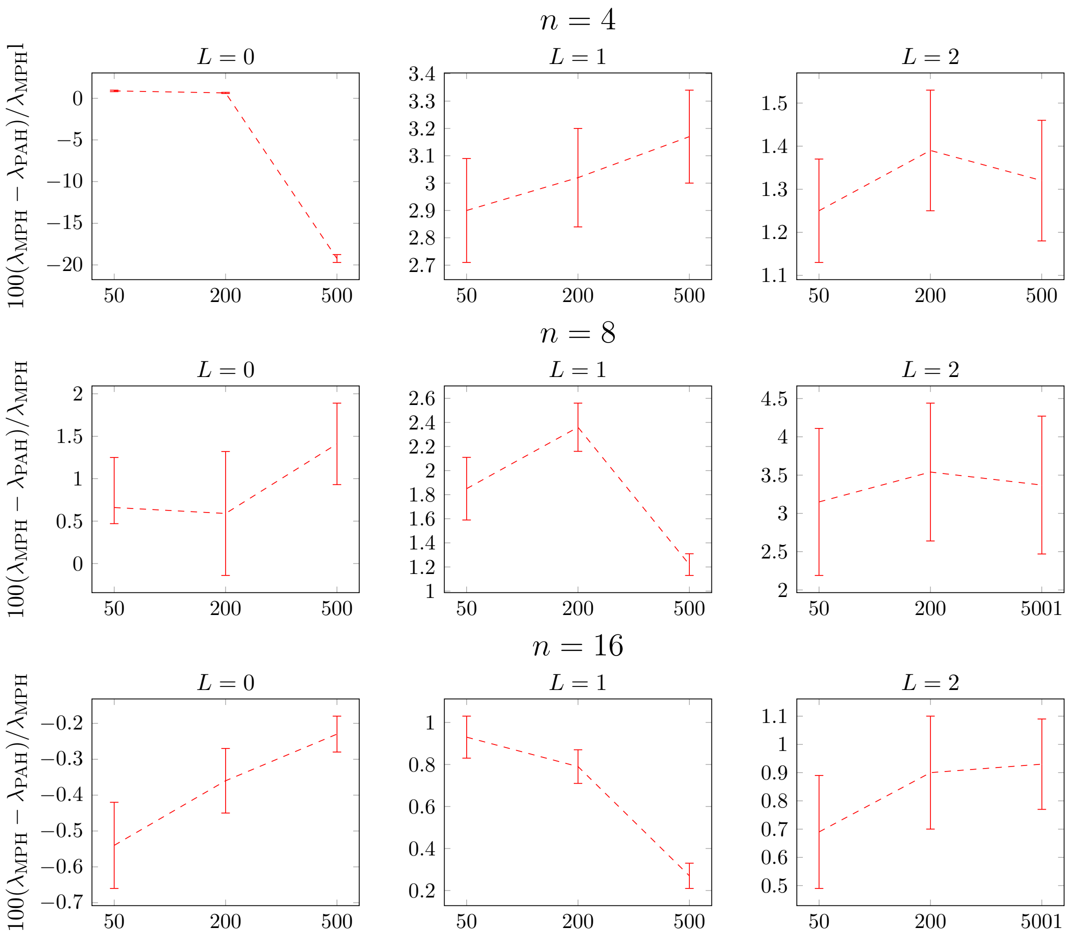
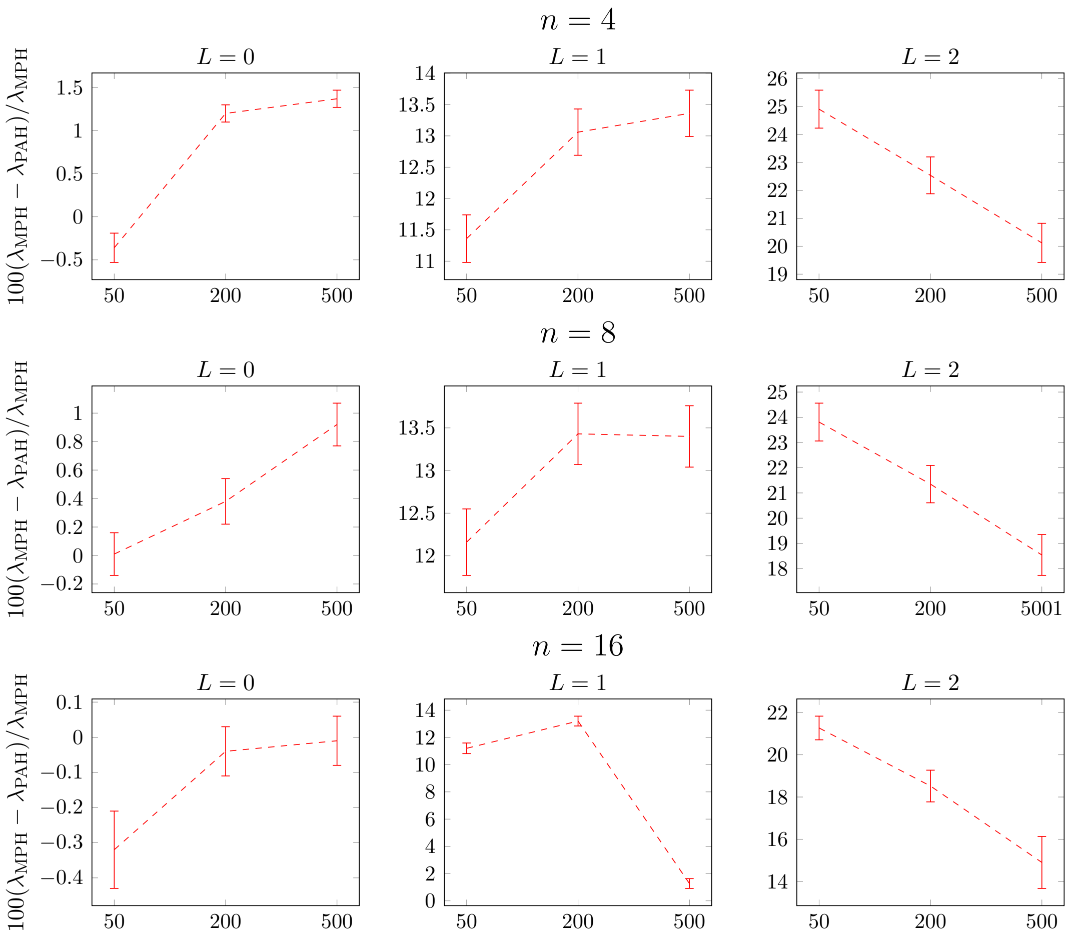


Figure 6.21: 95% confidence intervals for $\frac{100(\lambda_{\mathrm{MPH}}-\lambda_{\mathrm{PAH}})}{\lambda_{\mathrm{MPH}}}$ vs. R, for $N=4$, $R_P=7$, $R\in\{50,200,500\}$, $n\in\{4,8,16\}$, $L\in\{0,1,2\}$, $h=1$, $c=1$, $b=10$.



As L increases, the differences in the average costs between the initial base-stock levels obtained via the discretized myopic policy and the FPA-based local search method also increase. For non-zero lead times, there exists a larger difference in the average costs between the initial base-stock levels obtained via the discretized myopic policy and the FPA-based local search method, in comparison to the known transition and emission matrices case.

FPA vs Continuous Myopic Policy

We now compare the FPA-based local search method and the continuous myopic policy in terms of the average cost differences, when the transition and emission matrices are unknown. Figure 6.22, 6.23, and 6.24 illustrate the 95% confidence intervals of the percentage of the differences between the average costs of the FPA-based local search method and the continuous myopic policy. Note that λ_H and λ_{PAH} denote the average cost obtained by the continuous myopic policy and the FPA-based local search method, respectively, for the unknown transition matrix and demand distributions case:

Figure 6.22: 95% confidence intervals for $\frac{100(\lambda_{\text{H}}-\lambda_{\text{PAH}})}{\lambda_{\text{H}}}$ vs. R, for $N=2$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h=1$, $c=1$, $b=10$, 30 replications, 10,000 periods

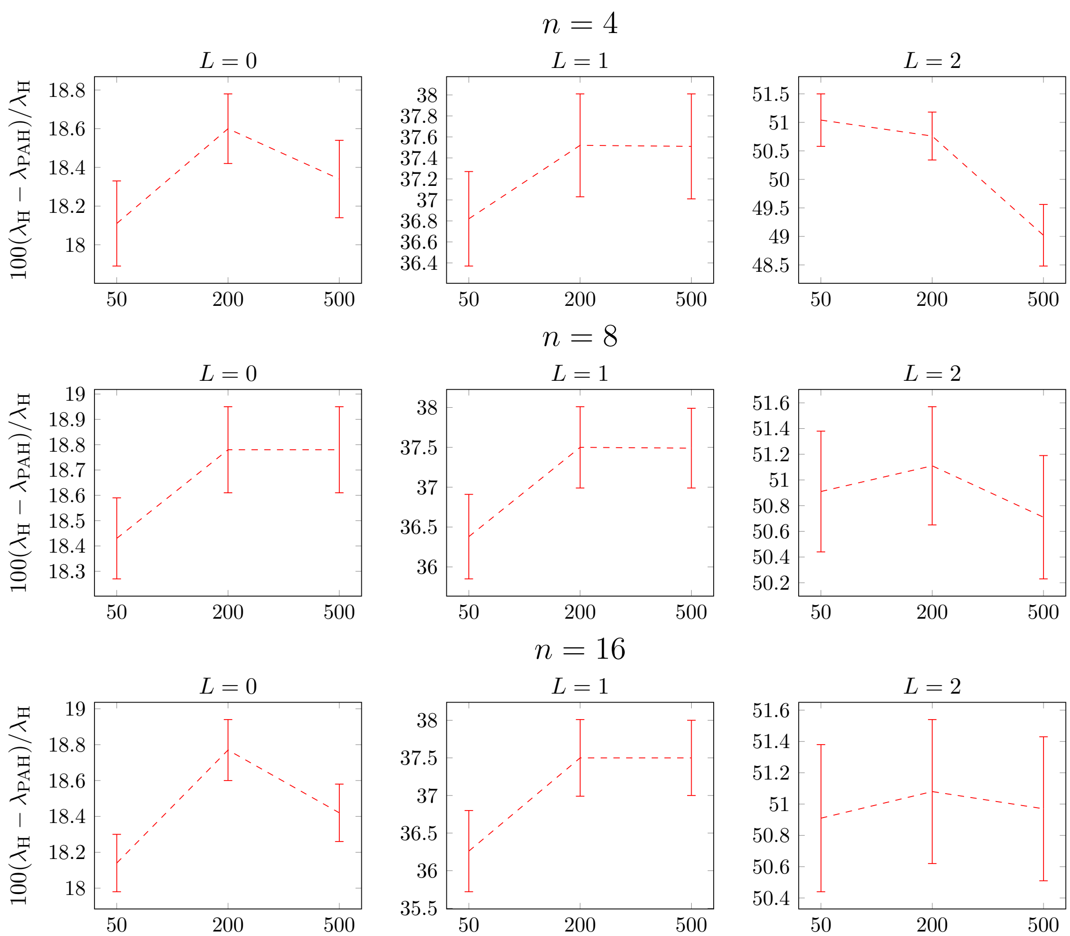


Figure 6.23: 95% confidence intervals for $\frac{100(\lambda_{\mathrm{H}}-\lambda_{\mathrm{PAH}})}{\lambda_{\mathrm{H}}}$ vs. R, for $N=3$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h=1$, $c=1$, $b=10$, 30 replications, 10,000 periods

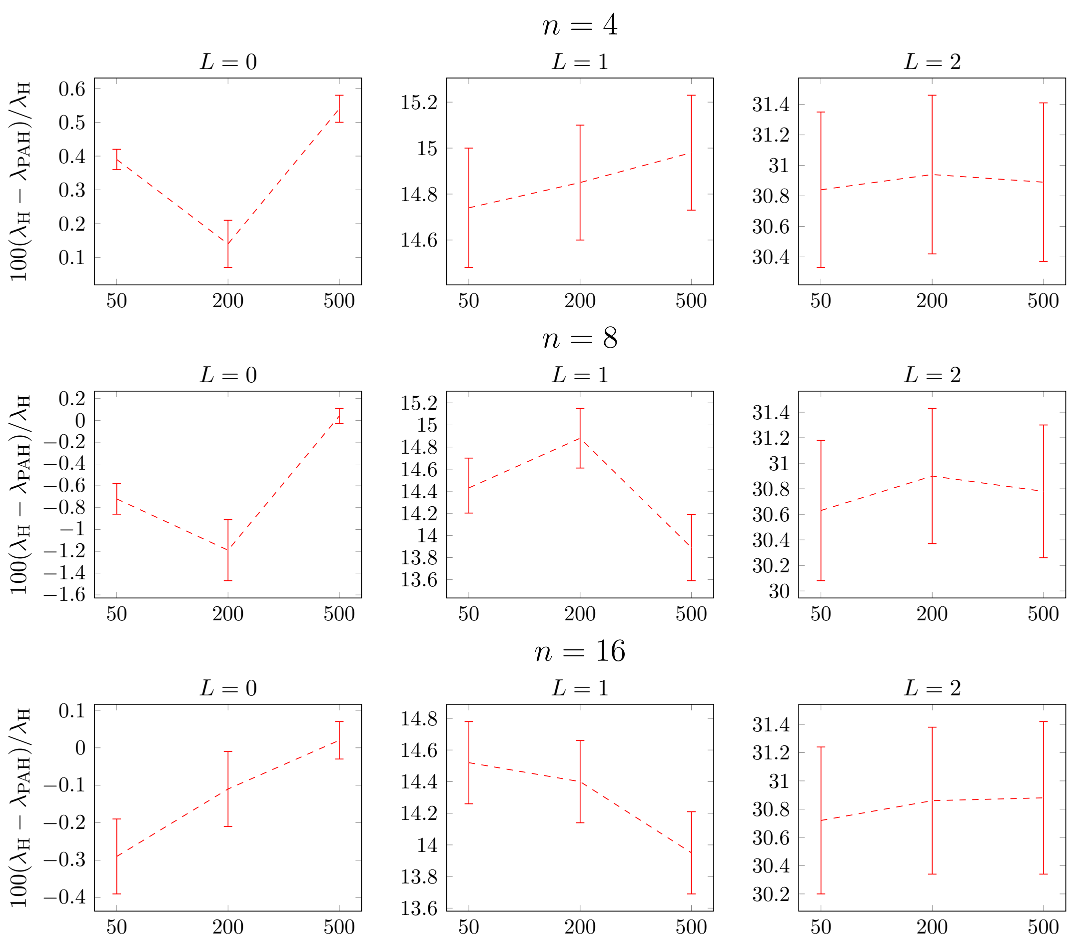
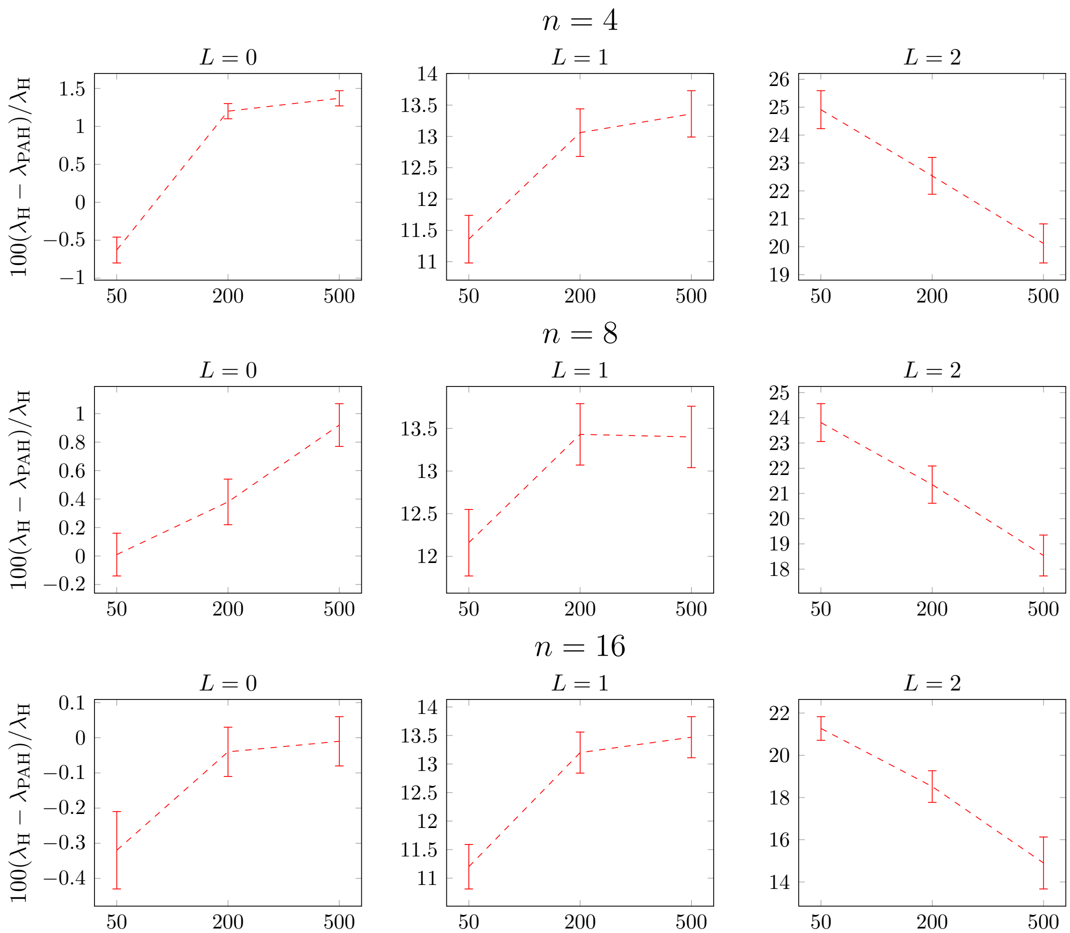


Figure 6.24: 95% confidence intervals for $\frac{100(\lambda_{\text{H}}-\lambda_{\text{PAH}})}{\lambda_{\text{H}}}$ vs. R, for $N=4$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h=1$, $c=1$, $b=10$, 30 replications, 10,000 periods



Figures 6.22, 6.23, and 6.24 show that when the transition and emission matrices are unknown, the FPA-based local search algorithm outperforms the continuous myopic policy in terms of the average cost for non-zero lead times. However, if the lead time is zero and $N \geq 3$, the FPA-based local search algorithm cannot perform as good as non-zero lead times. For some cases we consider, the FPA-based local search algorithm barely outperforms the continuous myopic policy, i.e., less than 1%. There are also some cases that the FPA-based local search algorithm shows slightly worse performance (at most approximately 1.3% worse), than the continuous myopic policy. For $N=2$, the FPA-based local search algorithm outperforms the continuous myopic policy for all cases that we consider.

6.2.3 FPA in Hidden vs. Known Transition and Emission Matrices

We now analyze the effect of having the information of transition and emission matrices. Recall that we denote by the average cost by λ_{PA} when the transition and emission matrix is known, and denote by λ_{PAH} when the transition and emission matrices are unknown. In Figures 6.25, 6.26, and 6.27, we analyze the effect of having transition and emission matrix information, via confidence intervals of the difference percentages of the average cost in known and unknown transition and emission matrix cases:

Figure 6.25: 95% confidence intervals for $\frac{100(\lambda_{\text{PAH}}-\lambda_{\text{PA}})}{\lambda_{\text{PA}}}$ vs. R, for $N=2$, $R\in\{50,200,500\}$, $n\in\{4,8,16\}$, $L\in\{0,1,2\}$, $h=1$, $c=1$, $b=10$.

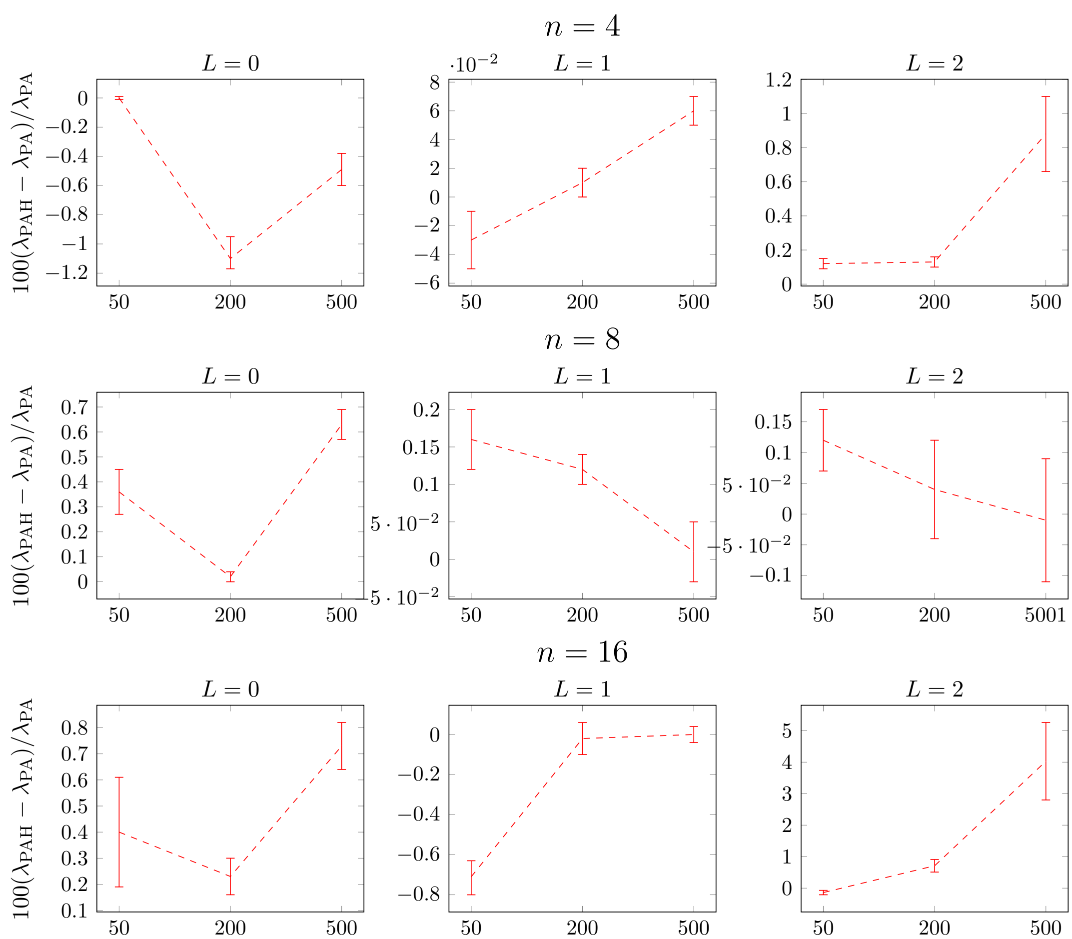


Figure 6.26: 95% confidence intervals for $\frac{100(\lambda_{\text{PAH}}-\lambda_{\text{PA}})}{\lambda_{\text{PA}}}$ vs. R, for $N = 3$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$.

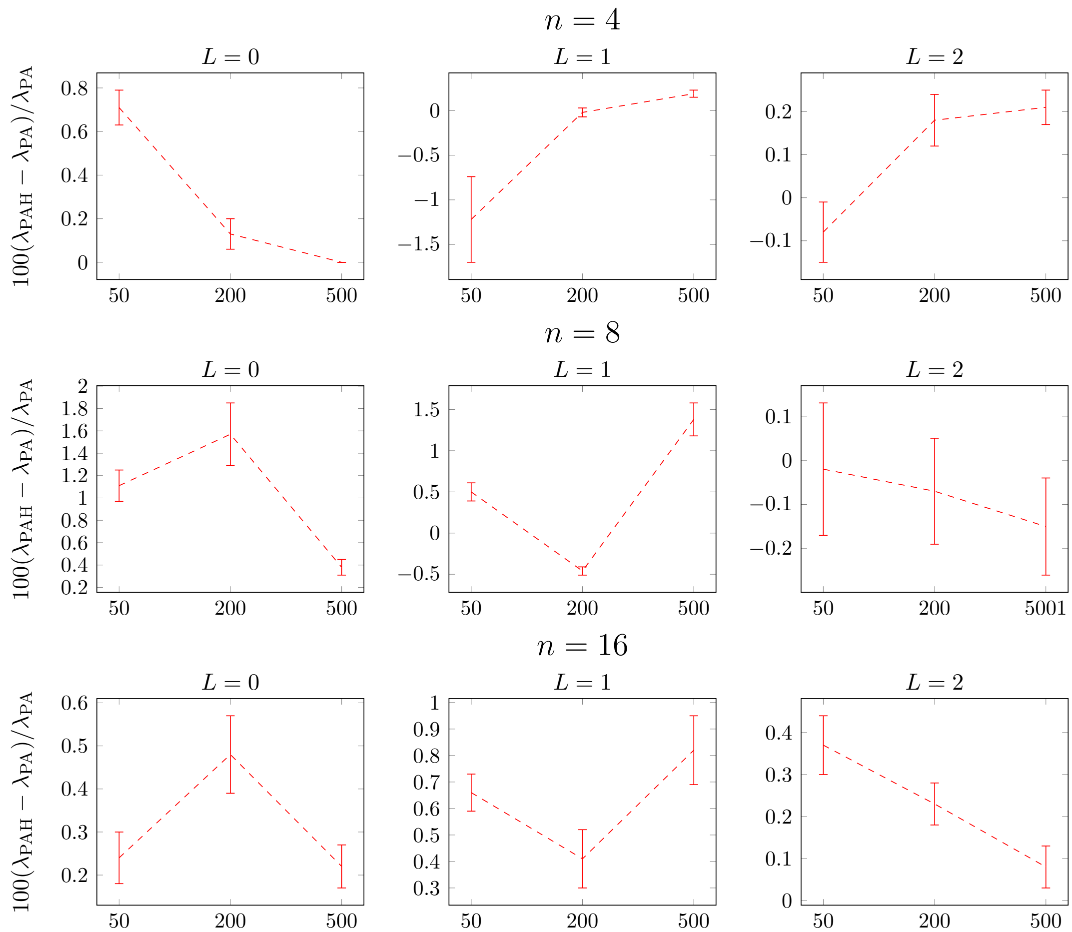
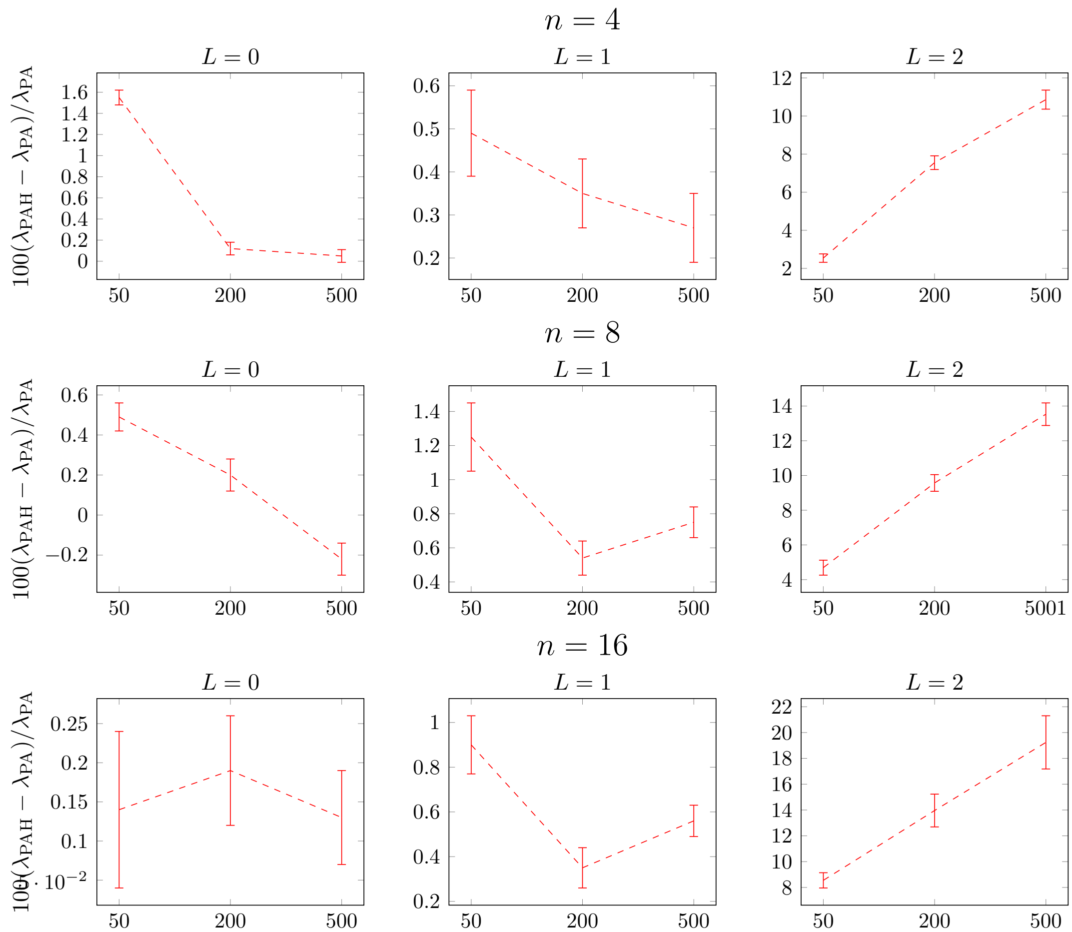


Figure 6.27: 95% confidence intervals for $\frac{100(\lambda_{\text{PAH}} - \lambda_{\text{PA}})}{\lambda_{\text{PA}}}$ vs. R, for $N = 4$, $R \in \{50, 200, 500\}$, $n \in \{4, 8, 16\}$, $L \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$.



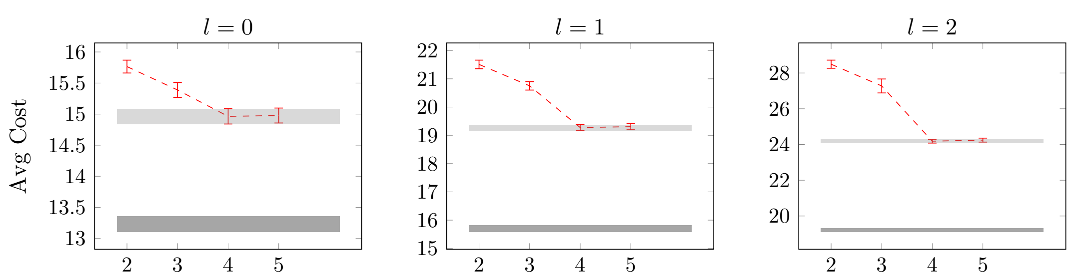
As seen from Figures 6.25, 6.27, and 6.27, when the transition matrix and demand distribution is unknown, the FPA-based local search still performs almost as good as in the known transition matrix and demand distributions case. In most cases, the performance of the method can be a little bit poorer in the unknown transition matrix and demand distribution case. The differences between the average costs obtained in both cases via the FPA-based local search method generally around 0-1 percent. However, there are some exceptions up to above 10% (i.e., $M = 4$, $n = 16$, $L = 2$).

6.2.4 Unknown Number of Demand States

Consider a given demand realization sequence, whose number of demand states is unknown, along with the transition and the emission matrix. Recall that when the number of demand states is unknown, we arbitrarily choose a number of demand states, $\tilde{N}$, and perform all calculations under the assumption that the estimated number of demand states is the true number of demand states.

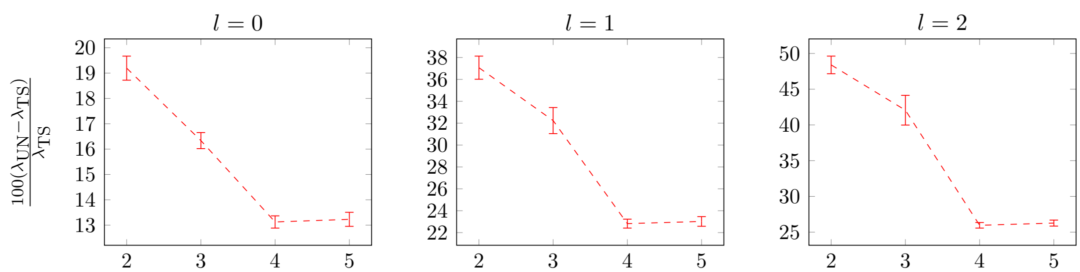
In our numerical experiment, we consider a demand sequence with 4 demand states. However, this is a hidden information. We estimate the number of states $\tilde{N}$ as 2, 3, 4, and 5, respectively. Then we apply the estimated number of demand states into the myopic policy. Figure 6.28 illustrates the estimated number of states in myopic policy, note that λ_{UN} and λ_{TS} represent the average costs between the cases in which no model parameters and also N is unknown, and everything including the true state is observable, respectively:

Figure 6.28: 95% confidence intervals for average cost vs. the number of demand distributions, in which the number of demand distributions are unknown, for update interval 7, maximum iterations 600, $l \in \{0, 1, 2\}$, $h = 1$, $c = 1,\ b = 10$. Light gray area is the confidence interval of the average cost in which the number of demand states is known and the emission and transition matrices are unknown, for $N = 4$. Dark gray area is the confidence interval of the average cost in which the true states, the emission and transition matrices and the number of demand distribution are known.



And Figure 6.29 illustrates the percentage of the difference of the average costs between the unknown number of demand states and the perfectly known demand states cases, which are denoted by λ_{UN} and λ_{TS}, respectively:

Figure 6.29: 95% confidence intervals for $\frac{100(\lambda_{\text{UN}}-\lambda_{\text{TS}})}{\lambda_{\text{TS}}}$ vs. the number of demand distributions, in which the number of demand distributions are unknown, for update interval 7, maximum iterations 600, $l \in \{0, 1, 2\}$, $h = 1$, $c = 1$, $b = 10$. Light gray area is the confidence interval of the average cost in which the number of demand states is known and the emission and transition matrices are unknown, for $N = 4$. Dark grey area is the confidence interval of the average cost in which the true states, the emission and transition matrices and the number of demand distribution is known.



As the estimated number of demand states $\tilde{N}$ approaches to the true number of demand states N, the average cost decrease until $\tilde{N} = N$ holds, then the confidence interval of the average cost begins to overlap with the unknown transition and emission matrices with known number of demand states case. The confidence interval of the average cost in this case is highlighted by a light gray area. When $\tilde{N} \geq N$, the confidence interval of the average cost remains overlapped with the light gray area.

Chapter 7

Conclusion

In this study, we consider a periodic-review inventory problem with discrete-valued Markov-modulated demand and partial information, where the demand distribution state process is characterized as a hidden Markov model with N different demand distribution states and the demand distribution state is only partially observed through demand realizations. Previous demand observations and the initial state belief vector yield a sufficient statistics for state belief. The problem is defined on a continuous state space. The inventory system operates under base-stock ordering policy and every state belief has its own base-stock level. There exist linear holding, shortage and ordering costs per item.

We propose a solution method based on finite perturbation analysis (FPA), that aims to find base-stock values which minimize the long-run average cost per period. We firstly discretize the continuous state space for various grid sizes, thus we obtain finite number of state beliefs and base-stock levels. Then, to find good starting points for the FPA-based local search algorithm, we use the discretized myopic policy.

The FPA-based local search algorithm updates the base-stock values during a single-replication simulation run, once in each update interval, by using the past inventory position, demand realization, order replenishment and state belief information. Note that the update interval must be selected as sufficiently large, such that it allows different state observations that ensure the neighboring base-stock vector with the lowest cost. In each update interval, the method calculates the effect of one-unit positive and negative changes in only one base-stock level, as well as calculating the average cost of the current base-stock levels. Then, it updates the base-stock levels with respect to the neighboring base-stock vector that yields the lowest cost. After some time, base-stock levels reach near-optimal values and starts to oscillate around the optimal values. After that point, the method finds near-optimal base-stock levels.

We implement the FPA-based local search method and construct a simulation model to calculate the average cost using MATLAB. We first consider the known transition and emission matrices case. We compare the average cost results of the initial base-stock levels obtained via the discretized myopic policy and the base-stock levels obtained via the FPA-based local search algorithm, for various number of demand states, grid sizes, and zero and positive lead times. We then compare the FPA-based local search algorithm with the continuous myopic policy, Viterbi algorithm and the sufficient statistic methods.

For the known transition and emission matrices case, the FPA-based local search method performs slightly better than the discretized myopic policy in terms of the average cost in most cases we consider. The FPA-based local search method approach to the performance of the continuous myopic policy so as to yield even less than 1% difference in the average cost. However, it fails to outperform the continuous myopic policy in general.

Later, we extend the FPA-based local search algorithm for unknown transition matrix and demand distributions. We compare the FPA-based local search method with the discretized and continuous myopic policies for this case. Then, we compare the results obtained via the FPA-based local search method for both the known and the unknown transition and emission matrices cases.

For the unknown transition and emission matrices case, we cannot conclude that the FPA-based local search method outperforms the discretized myopic policy when the lead time is zero. For non-zero lead times, the FPA-based local search method outperforms the discretized myopic policy in terms of the average cost. The FPA-based method also outperforms the continuous myopic policy for non-zero lead times, and both methods have a similar performance when the lead time is zero. For both known and unknown transition and emission matrices cases, the percentages of differences results obtained via the FPA-based local search method is not significant, below 1% in most cases, except when $N = 4$ and $L = 2$.

We then analyze the unknown number of demand states case and in terms of its effects on the average cost. We observe the relation between the estimated number of demand states and the average cost, We compare the average cost results for the estimated number of demand states, the unknown transition and emission matrices, and known actual states, transition and emission matrices cases. As the estimated number of demand states increases and gets closer to the actual number of demand states, the average cost tend to decrease. As the estimated number of demand states is greater than or equal to the actual number of demand states, the 95% confidence intervals of the average costs for the unknown number of demand states and the unknown transition and emission matrices cases overlap.

Our algorithm updates a single base-stock level by 1 unit in each update interval, in other words it has a step size 1. One possible future direction would be developing an FPA-based method that estimates the effects of differences of finite values larger than 1. Increased step size may speed up the method in terms of the number of periods required for convergence, considering the cases with large and sudden changes in base-stock levels. An extension to our study may be considering to update more than 1 base-stock levels for each update interval. A future direction may be to combine the myopic policy and the FPA-based local search algorithm. Developing a heuristic in order to estimate the number of demand states throughout the realized demand values may be another future direction.

Bibliography

[1] J. S. Song and P. Zipkin, "Inventory control in a fluctuating demand environment," *Operations Research*, vol. 41, no. 2, pp. 351 – 370, 1993.

[2] S. P. Sethi and F. Cheng, "Optimality of (s, s) policies in inventory models with markovian demand," *IEEE Transactions on Medical Imaging*, vol. 45, no. 6, pp. 931 – 939, 1997.

[3] D. Beyer and S. P. Sethi, "Average cost optimality in inventory models with markovian demands," *IEEE Transactions on Medical Imaging*, vol. 92, no. 3, pp. 497 – 526, 1997.

[4] J. T. Treharne and C. R. Sox, "Adaptive inventory control for nonstationary demand and partial information," *Management Science*, vol. 48, no. 5, pp. 607 – 624, 2002.

[5] K. Arifoglu and S. Ozekici, "Optimal policies for inventory systems with finite capacity and partially observed markov-modulated demand and supply processes," *European Journal of Operational Research*, vol. 203, no. 3, pp. 421 – 438, 2010.

[6] E. Bayraktar and M. Ludkovski, "Inventory management with partially observed nonstationary demand," *Annals of Operations Research*, vol. 176, no. 1, pp. 7 – 39, 2010.

[7] A. Viterbi, "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Transactions on Information Theory*, vol. 13, no. 2, pp. 260 – 269, 1967.

[8] L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state markov chains," *The Annals of Mathematical Statistics*, vol. 37, no. 6, pp. 1554 – 1563, 1966.

[9] L. E. Baum and J. A. Eagon, "An inequality with applications to statistical estimation for probabilistic functions of markov processes and to a model for ecology," *Bulletin of the American Mathematical Society*, vol. 73, no. 4, pp. 360 – 363, 1967.

[10] L. E. Baum and G. Sell, "Growth transformations for functions on manifolds," *Pacific Journal of Mathematics*, vol. 27, no. 2, pp. 211 – 227, 1968.

[11] G. S. L. E. Baum, T. Petrie and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains," *The annals of mathematical statistics*, vol. 41, no. 1, pp. 164 – 171, 1970.

[12] L. E. Baum, "An inequality and associated maximization technique in statistical estimation of probabilistic functions of a markov process," *Inequalities*, vol. 3, pp. 1 – 8, 1972.

[13] T. Vercauteren and A. L. Toledo, "Batch and sequential bayesian estimators of the number of active terminals in an ieee 802.11 network," *IEEE Transactions on Signal Processing*, vol. 55, no. 2, pp. 437 – 450, 2007.

[14] M. Y. T. Iki, M. Horiguchi and M. Kurano, "A learning algorithm for communcating markov decision processes with unknown transition matrices," *Bulletin of Informatics and Cybernetics*, vol. 39, 2007.

[15] M. A. E. Y. C. Ho and T. T. Chien, "A gradient technique for general buffer storage design in a production line," *International journal of production research*, vol. 17, no. 6, pp. 557 – 580, 1979.

[16] Y. C. Ho and X. R. Cao, "Perturbation analysis and optimization of queueing networks," *Journal of Optimization Theory and Applications*, vol. 40, no. 4, pp. 559 – 582, 1983.

[17] X. R. C. Y. C. Ho and C. Cassandras, "Infinitesimal and finite perturbation analysis for queueing networks," *Automatica*, vol. 19, no. 4, pp. 439 – 445, 1983.

[18] X. R. Cao, "Convergence of parameter sensitivity estimates in a stochastic environment," *IEEE Transactions on Automatie Control AC-30*, pp. 834 – 843, 1985.

[19] X. C. Y. C. Ho and C. Cassandras, "Perturbation analysis gives strongly consistent sensitivity estimates for the m/gil queue," *Management Science*, vol. 34, no. 1, pp. 39 – 64, 1988.

[20] Y. C. Ho and X. R. Cao, "Performance sensitivity to routing changes in queueing networks and flexible manufacturing systems using perturbation analysis," *IEEE Journal on Robotics and Automation*, vol. 1, pp. 165 – 172, 1985.

[21] P. Glasserman, "Equivalence methods in the perturbation analysis of queueing networks," *Division of Applied Science*, 1988a.

[22] P. Vakili, "Three topics on perturbation analysis ofdiscrete-event dynamic systems," 1989.

[23] X. R. Cao, "Realization probability in closed queueing networks and its application," *Advances in Applied Probability*, vol. 19, pp. 708 – 738, 1987a.

[24] P. L'Ecuyer, "A unified view of the ipa, sf, and lr gradient estimation techniques," *Management Science*, vol. 36, no. 11, pp. 1364 – 1383, 1990.

[25] P. Glasserman, "Structural condition for perturbation analysis derivative estimation: Finite time perfonnance indices," *Journal 0/ the ACM*, 1990b.

[26] J. Q. Hu, "Consistency of infinitesimal perturbation analysis estimators with load-dependent rates," *Queueing Systems*, 1991.

[27] P. Glasserman and D. Yao, "Algebraic structure of some stochastic discrete event systems with applications," *Journal of Discrete Event Dynamic Systems*, vol. 1, no. 1, 1991.

[28] S. Bashyam and M. C. Fu, "Application of perturbation analysis to (s, s) inventory systems," *Simulation Conference*, 1991.

[29] S. Bashyam and M. C. Fu, "Application of perturbation analysis to a class of periodic review (s, s) inventory systems," *Naval Research Logistics (NRL)*, vol. 41, no. 1, pp. 47 – 80, 1994.

[30] M. C. Fu and K. J. Healy, "Simulation optimization of (s, s) inventory systems," *Proceedings of the 24th conference on Winter simulation. ACM*, 1992.

[31] K. J. Healy and L. W. Schruben, "Retrospective simulation response optimization," *Proceedings of the 23rd conference on Winter simulation. IEEE Computer Society*, 1991.

[32] M. C. Fu and K. J. Healy, "Techniques for optimization via simulation: an experimental study on an (s, s) inventory system," *IIE transactions*, vol. 29, no. 3, pp. 191 – 199, 1997.

[33] M. C. Fu and K. J. Healy, "Sample path derivatives for (s, s) inventory systems," *Operations Research*, vol. 42, no. 2, pp. 351 – 364, 1994.

[34] F. J. Vazquez-Abad and M. Cepeda-Juneman, "The phantom spa method: An inventory problem revisited," *Proceedings of the 1999 Winter Simulation Conference*, pp. 1664 – 1669, 1999.

[35] S. Bashyam and M. C. Fu, "Optimization of (s, s) inventory systems with random lead times and a service level constraint," *Management Science*, vol. 44, no. 12, pp. 243 – 256, 1998.

[36] H. Zhang and M. C. Fu, "Sample path derivatives for (s, s) inventory systems," 2005.

[37] D. M. L. Karim and H. Chen, "Parameter perturbation analysis through stochastic petri nets: Application to an inventory system," 2010.

[38] D. M. L. Karim and H. Chen, "Application of perturbation analysis to multiproduct capacitated production-inventory control," *Proceedings of theAmerican Control Conference*, pp. 2170 – 2174, 1995.

[39] P. Glasserman and S. Tayur, "Sensitivity analysis for base-stock levels in multiechelon production-inventory systems," *Management Science*, vol. 41, no. 2, pp. 263 – 281, 1995.

[40] R. Anupindi and S. Tayur, "Managing stochastic multiproduct systems: model, measures, and analysis," *INFORMS*, vol. 46, no. 3, pp. 98 – 111, 1997.

[41] R. Kapuscinski and S. Tayur, "A capacitated production-inventory model with periodic demand," *Operations Research*, vol. 46, no. 6, pp. 899 – 911, 1998.

[42] C. G. C. I. C. Paschalidis, Yong Liu and P. Zhang, "Threshold-based control for make-to-stock models: A synergy between large deviations and perturbation analysis," *Proceedings of the 40th IEEE Conference on Decision and Control*, pp. 4523 – 4528, 2001.

[43] C. G. C. I. C. Paschalidis, Yong Liu and C. Panayiotou, "Inventory control for supply chains with service level constraints: A synergy between large deviations and perturbation analysis," *Annals of Operations Research*, vol. 126, pp. 231 – 258, 2004.

[44] Y. Zhao and B. Melamed, "Make-to-stock systems with backorders: Ipa gradients," *Proceedings of the 2004 Winter Simulation Conference*, 2004.

[45] Y. Zhao and B. Melamed, "Ipa derivatives for make-to-stock production-inventory systems with backorders," *Methodol Comput Appl Probab*, vol. 8, pp. 191 – 222, 2006.

[46] Y. Zhao and B. Melamed, "Inventory control for supply chains with service level constraints: A synergy between large deviations and perturbation analysis," *IEEE TRANSACTIONS ON AUTOMATIC CONTROL*, vol. 52, no. 8, pp. 1491 – 1495, 2007.

[47] Y. Z. Y. Fan, B. Melamed and Y. Wardi, "Ipa derivatives for make-to-stock production-inventory systems with backorders under the (r,r) policy," *Methodol Comput Appl Probab*, vol. 11, pp. 159 – 179, 2009.

[48] Y. Z. B. Melamed, Y. Fan and Y. Wardi, “Ipa derivatives for a discrete model of make-to-stock production-inventory systems with backorders,” *Annals of Operations Research*, vol. 181, pp. 1 – 19, 2010.

[49] E. Y. D. Ozdemir and Y. T. Herer, “Multi-location transshipment problem with capacitated transportation,” *European Journal of Operational R*, vol. 175, pp. 602 – 621, 2005.

[50] L. H. L. E. P. Chew and Y. L. Lau, “Component allocation and ordering policy in a multi-component, multi-product assembled to stock system,” *OR Spectrum*, vol. 32, pp. 293 – 317, 2010.

[51] W. W. J. D. Schwartz and D. E. Rivera, “Simulation-based optimization of process control policies for inventory management in supply chains,” *Automatica*, vol. 42, pp. 1311 – 1320, 2006.

[52] E. E. Ozguven and K. Ozbay, “A secure and efficient inventory management system for disasters,” *Transportation Research Part C*, pp. 171 – 196, 2011.

[53] H. Z. T. Sadok and R. Nidhal, “Ipa derivatives for a discrete model of make-to-stock production-inventory systems with backorders,” *7th IFAC Conference on Manufacturing Modelling, Management, and Control*, 2013.

[54] Z. H. S. Ayed and N. Rezg, “Fpa method for optimal production planning under availability/degradation machine and subcontracting constraint,” *International Journal of Production Researchh*, vol. 55, no. 8, pp. 2135 – 2148, 2017.

[55] J. A. Nelder and R. Mead, “A simplex method for function minimization,” *The computer journal*, vol. 7, no. 4, pp. 308 – 313, 1965.

[56] W. J. Stevenson, “Operations management, 10th ed.,” pp. 581 – 582, 2009.

[57] D. F. P. E. A. Silver and R. Peterson, “Inventory management and production planning and scheduling, 3rd ed.,” pp. 385 – 392, 1998.

Appendix A

Tables

Table A.1: FPA-based local search algorithm calculation times for available HMM parameter information, in seconds, with the corresponding demand distributions and transition matrices for each N, $c = 1,\ h = 1,\ b = 10$.

N	n	$L = 0$			$L = 1$			$L = 2$		
		$R = 50$	$R = 200$	$R = 500$	$R = 50$	$R = 200$	$R = 500$	$R = 50$	$R = 200$	$R = 500$
2	4	0.4063	0.0156	0.0313	0.2344	0.0313	0.0313	0.1094	0.0313	0.0313
	8	0.0469	0.0625	0.0156	0.0469	0.0469	0.0313	0.1250	0.0469	0.0469
	16	0.0313	0.0313	0.0313	0.0469	0.0156	0.0313	0.0781	0.0938	0.0781
3	4	0.2188	0.0625	0.0313	0.2656	0.0469	0.0469	0.2031	0.0781	0.0781
	8	0.1250	0.0625	0.0469	0.0938	0.0625	0.0625	0.2031	0.2188	0.2031
	16	0.0781	0.0469	0.0469	0.1406	0.1406	0.1094	0.5938	0.6094	0.5781
4	4	0.3438	0.0625	0.0313	0.1719	0.0625	0.1094	0.2969	0.2031	0.2188
	8	0.1094	0.1875	0.0781	0.1406	0.1250	0.1406	0.8594	0.8594	0.8125
	16	0.3438	0.2188	0.2031	0.6406	0.5938	0.5781	4.6094	4.6094	4.5625

Table A.2: 95% confidence intervals for the average costs, initial base-stock levels found via the discretized myopic policy vs FPA, $N = 2$, $c = 1$, $h = 1$ $b = 10$.

n	95% CI	L = 0				L = 1				L = 2			
		Initial	R = 50	R = 200	R = 500	Initial	R = 50	R = 200	R = 500	Initial	R = 50	R = 200	R = 500
4	CI Upper	19.7319	19.3795	19.5404	19.4143	27.9744	27.9941	27.4961	27.4922	36.8484	35.6841	35.5365	35.5581
	Avg	19.6244	19.3251	19.4584	19.3397	27.8492	27.9399	27.4362	27.4291	36.6838	35.5700	35.4220	35.4450
	CI Lower	19.5169	19.2707	19.3765	19.2651	27.7240	27.8857	27.3763	27.3660	36.5192	35.4558	35.3075	35.3319
8	CI Upper	19.2005	19.3795	19.2960	19.2546	27.5733	27.9941	27.4961	27.4961	35.8892	35.6841	35.5852	35.6841
	Avg	19.1785	19.3251	19.2419	19.2074	27.5142	27.9399	27.4362	27.4362	35.7636	35.5700	35.4726	35.5700
	CI Lower	19.1565	19.2707	19.1877	19.1601	27.4551	27.8857	27.3763	27.3763	35.6380	35.4558	35.3601	35.4558
16	CI Upper	19.2005	19.3795	19.2960	19.2546	27.5733	27.9941	27.4961	27.4961	35.8892	35.6841	35.5852	35.6841
	Avg	19.1785	19.3251	19.2419	19.2074	27.5142	27.9399	27.4362	27.4362	35.7636	35.5700	35.4726	35.5700
	CI Lower	19.1565	19.2707	19.1877	19.1601	27.4551	27.8857	27.3763	27.3763	35.6380	35.4558	35.3601	35.4558

Table A.3: 95% confidence intervals for the average costs, initial base-stock levels found via the discretized myopic policy vs FPA, $N = 3$, $c = 1$, $h = 1$ $b = 10$.

n	95% CI	L = 0				L = 1				L = 2			
		Initial	R = 50	R = 200	R = 500	Initial	R = 50	R = 200	R = 500	Initial	R = 50	R = 200	R = 500
4	CI Upper	16.2066	16.1158	16.1029	16.0565	22.2963	21.6986	21.6459	21.5666	28.3020	28.0628	27.9065	27.9060
	Avg	16.1163	16.0291	16.0180	15.9759	22.1713	21.5955	21.5543	21.4750	28.1644	27.9758	27.7843	27.7943
	CI Lower	16.0261	15.9424	15.9331	15.8953	22.0463	21.4627	21.5111	21.3835	28.0268	27.8887	27.6620	27.6826
8	CI Upper	16.4357	16.0799	16.0849	16.0762	22.1245	21.7701	21.7158	21.5787	29.2195	27.9351	21.9562	28.0270
	Avg	16.3609	15.9981	16.0026	15.9950	22.0336	21.6662	21.6411	21.4960	29.1348	27.8441	27.8671	27.9380
	CI Lower	16.2861	15.9981	15.9203	15.9137	21.9427	21.5623	21.5665	21.4134	29.0501	27.7531	27.7780	27.8489
16	CI Upper	16.0778	16.1136	16.0899	16.1024	21.9718	21.7051	21.6611	21.6967	28.1267	28.0317	27.9164	27.9343
	Avg	16.0021	16.0240	16.0035	16.0238	21.8767	21.6062	21.5728	21.5982	28.0528	27.9467	27.8003	27.8322
	CI Lower	15.9265	15.9345	15.9172	15.9451	21.7817	21.5073	21.4845	21.4996	27.9790	27.8616	27.6842	27.7300

Table A.4: 95% confidence intervals for the average costs, initial base-stock levels found via the discretized myopic policy vs FPA, $N = 4$, $c = 1$, $h = 1$ $b = 10$.

n	95% CI	L = 0				L = 1				L = 2			
		Initial	R = 50	R = 200	R = 500	Initial	R = 50	R = 200	R = 500	Initial	R = 50	R = 200	R = 500
4	CI Upper	15.1439	15.2063	15.1416	15.1278	19.7729	19.5326	19.4857	19.4380	24.5584	24.4423	24.3034	24.3089
	Avg	15.0208	15.0873	15.0213	15.0061	19.6384	19.4186	19.3717	19.3196	24.4241	24.3258	24.1848	24.1962
	CI Lower	14.8976	14.9682	14.9010	14.8845	19.5040	19.3046	19.2577	19.2012	24.2898	24.2093	24.0662	24.0836
8	CI Upper	15.1710	15.1418	15.1944	15.1763	19.6195	19.5724	19.4927	19.4598	24.7564	24.4766	24.2559	24.3089
	Avg	15.0493	15.0160	15.0702	15.0532	19.5003	19.4537	19.3769	19.3416	24.6608	24.3625	24.1505	24.2495
	CI Lower	14.9276	14.8902	14.9459	14.9300	19.3811	19.3351	19.2611	19.2233	24.5653	24.2483	24.0452	24.0373
16	CI Upper	15.0870	15.1372	15.1302	15.1358	19.7729	19.5597	19.4867	19.3861	24.3116	24.5520	24.3375	24.2975
	Avg	14.9687	15.0103	15.0148	15.0148	19.4364	19.3737	19.2725	19.2726	24.4371	24.2194	24.1812	24.1813
	CI Lower	14.8505	14.8942	14.8903	14.8939	19.5040	19.3130	19.2606	19.1590	24.2898	24.3223	24.1012	24.0650

Table A.5: 95% confidence intervals for the average costs, continuous myopic policy vs FPA, $N = 2$, $c = 1$, $h = 1$ $b = 10$.

n	95% CI	L = 0				L = 1				L = 2			
		Myopic	R = 50	R = 200	R = 500	Myopic	R = 50	R = 200	R = 500	Myopic	R = 50	R = 200	R = 500
4	CI Upper	19.1842	19.3795	19.5404	19.4143	27.5008	27.9941	27.4961	27.4922	35.5389	35.6841	35.5365	35.5581
	Avg	19.1535	19.3251	19.4584	19.3397	27.5142	27.9399	27.4079	27.4291	35.4188	35.5700	35.4220	35.4450
	CI Lower	19.1228	19.2707	19.3765	19.2651	27.4551	27.8857	27.3150	27.3660	35.2988	35.4558	35.3075	35.3319
8	CI Upper	19.1842	19.3795	19.2960	19.2546	27.5008	27.9941	27.4961	27.4961	35.5389	35.6841	35.5852	35.6841
	Avg	19.1535	19.3251	19.2419	19.2074	27.5142	27.9399	27.4079	27.4362	35.4188	35.5700	35.4726	35.5700
	CI Lower	19.1228	19.2707	19.1877	19.1601	27.4551	27.8857	27.3150	27.3763	35.2988	35.4558	35.3601	35.4558
16	CI Upper	19.1842	19.3795	19.2960	19.2546	27.5008	27.9941	27.4961	27.4961	35.5389	35.6841	35.5852	35.6841
	Avg	19.1535	19.3251	19.2419	19.2074	27.5142	27.9399	27.4079	27.4362	35.4188	35.5700	35.4726	35.5700
	CI Lower	19.1228	19.2707	19.1877	19.1601	27.4551	27.8857	27.3150	27.3763	35.2988	35.4558	35.3601	35.4558

Table A.6: 95% confidence intervals for the average costs, continuous myopic policy vs FPA, $N = 3$, $c = 1$, $h = 1$ $b = 10$.

n	95% CI	L = 0				L = 1				L = 2			
		Myopic	R = 50	R = 200	R = 500	Myopic	R = 50	R = 200	R = 500	Myopic	R = 50	R = 200	R = 500
4	CI Upper	16.0462	16.1158	16.1029	16.0565	21.5780	21.6986	21.6459	21.5666	27.9897	28.0628	27.9065	27.9060
	Avg	15.9664	16.0291	16.0180	15.9759	21.4996	21.5955	21.5543	21.4750	27.9024	27.9758	27.7843	27.7943
	CI Lower	15.8862	15.9424	15.9331	15.8953	21.4213	21.4627	21.5111	21.3835	27.8151	27.8887	27.6620	27.6826
8	CI Upper	16.0462	16.0799	16.0849	16.0762	21.5780	21.7701	21.7158	21.5787	27.9897	27.9351	21.9562	28.0270
	Avg	15.9664	15.9981	16.0026	15.9950	21.4996	21.6662	21.6411	21.4960	27.9024	27.8441	27.8671	27.9380
	CI Lower	15.8862	15.9981	15.9203	15.9137	21.4213	21.5623	21.5665	21.4134	27.8151	27.7531	27.7780	27.8489
16	CI Upper	16.0462	16.1136	16.0899	16.1024	21.5780	21.7051	21.6611	21.6967	27.9897	28.0317	27.9164	27.9343
	Avg	15.9664	16.0240	16.0035	16.0238	21.4996	21.6062	21.5728	21.5982	27.9024	27.9467	27.8003	27.8322
	CI Lower	15.8862	15.9345	15.9172	15.9451	21.4213	21.5073	21.4845	21.4996	27.8151	27.8616	27.6842	27.7300

Table A.7: 95% confidence intervals for the average costs, continuous myopic policy vs FPA, $N = 4$, $c = 1$, $h = 1$ $b = 10$.

n	95% CI	L = 0				L = 1				L = 2			
		Myopic	R = 50	R = 200	R = 500	Myopic	R = 50	R = 200	R = 500	Myopic	R = 50	R = 200	R = 500
4	CI Upper	15.0549	15.2063	15.1416	15.1278	19.3276	19.5326	19.4857	19.4380	24.1924	24.4423	24.3034	24.3089
	Avg	14.9338	15.0873	15.0213	15.0061	19.2194	19.4186	19.3717	19.3196	24.0919	24.3258	24.1848	24.1962
	CI Lower	14.8126	14.9682	14.9010	14.8845	19.1112	19.3046	19.2577	19.2012	23.9913	24.2093	24.0662	24.0836
8	CI Upper	15.0549	15.1418	15.1944	15.1763	19.3276	19.5724	19.4927	19.4598	24.1624	24.4766	24.2559	24.3089
	Avg	14.9338	15.0160	15.0702	15.0532	19.2194	19.4537	19.3769	19.3416	24.0919	24.3625	24.1505	24.2495
	CI Lower	14.8126	14.8902	14.9459	14.9300	19.1112	19.3351	19.2611	19.2233	23.9913	24.2483	24.0452	24.0373
16	CI Upper	15.0549	15.1372	15.1302	15.1358	19.3276	19.5597	19.4867	19.3861	24.1924	24.5520	24.3375	24.2975
	Avg	14.9687	15.0103	15.0148	15.0148	19.2194	19.3737	19.2725	19.2726	24.0919	24.2194	24.1812	24.1813
	CI Lower	14.8126	14.8942	14.8903	14.8939	19.1112	19.3130	19.2606	19.1590	23.9913	24.3223	24.1012	24.0650

Table A.8: Comparison of 95% confidence intervals for the average costs obtained by Viterbi algorithm, sufficient statistic method, and FPA, $N = 3$, $c = 1$, $h = 1$ $b = 10$.

Algorithm	95% CI	L = 0				L = 1				L = 2			
		Initial	R=50	R=200	R=500	Initial	R=50	R=200	R=500	Initial	R=50	R=200	R=500
Viterbi	CI Upper	16.9477	16.9992	16.2179	16.2179	23.0303	23.5999	22.3210	22.3210	33.4783	32.2123	32.2123	32.2123
	Avg	16.8444	16.8989	16.1297	16.1297	22.8949	23.4494	22.2002	22.2002	33.3877	32.1361	32.1361	32.1361
	CI Lower	16.7412	16.7986	16.0415	16.0415	22.7596	23.2990	22.0793	22.0793	33.2972	32.0600	32.0600	32.0600
Suff. Stat.	CI Upper	16.9497	17.0012	16.2200	16.2200	23.0350	23.6051	22.3257	22.3257	33.4804	32.2144	32.2144	32.2144
	Avg	16.8463	16.9007	16.1316	16.1316	22.8991	23.4541	22.2043	22.2043	33.3897	32.1382	32.1382	32.1382
	CI Lower	16.7428	16.8002	16.0432	16.0432	22.7632	23.3031	22.0830	22.0830	33.2989	32.0619	32.0619	32.0619

Table A.9: Average cost results and 95% confidence intervals for the myopic policy, max iterations 600 and 1200 for each $R_P \in \{7, 14\}$, respectively, $N \in \{2, 3, 4\}$ $c = 1,\ h = 1$ $b = 10$.

N	95% CI	$L = 0$			$L = 1$			$L = 2$		
		Known	$R_P = 7$	$R_P = 14$	Known	$R_P = 7$	$R_P = 14$	Known	$R_P = 7$	$R_P = 14$
2	CI Upper	19.1842	23.7921	23.7909	27.5008	44.2985	44.8363	35.5389	73.0556	74.8464
	Average	19.1535	23.6944	23.6934	27.4079	43.9245	44.4453	35.4188	72.5777	74.3392
	CI Lower	19.1228	23.5967	23.5959	27.3150	43.5506	44.0542	35.2988	72.0998	73.8320
3	CI Upper	16.0462	16.1415	16.1411	21.5780	25.4319	25.6517	27.9897	40.6756	41.5198
	Average	15.9664	16.0620	16.0619	21.4996	25.3074	25.5259	27.9024	40.3190	41.1477
	CI Lower	15.8866	15.9825	15.9827	21.4213	25.1829	25.4002	27.8151	39.9624	40.7757
4	CI Upper	15.0549	15.1072	15.1163	19.3276	22.0328	22.2121	24.1924	32.7779	33.3666
	Average	14.9338	14.9878	14.9940	19.2194	21.9362	22.1059	24.0919	32.4912	33.0600
	CI Lower	14.8126	14.8685	14.8718	19.1112	21.8395	21.9996	23.9913	32.2046	32.7535

Table A.10: 95% confidence intervals for the average costs, initial base-stock levels found via the disretized myopic policy vs FPA, for unknown transition matrix and demand distribution case, $N = 2$, $R_P = 7$, $c = 1$, $h = 1$ $b = 10$

n	95% CI	L = 0				L = 1				L = 2			
		Initial	R = 50	R = 200	R = 500	Initial	R = 50	R = 200	R = 500	Initial	R = 50	R = 200	R = 500
	CI Upper	23.6897	19.4650	19.3017	19.3846	46.0703	28.0419	27.5006	27.5011	76.6743	35.7256	35.5998	35.6774
4	Avg	23.5999	19.3955	19.2463	19.3291	45.8318	27.9851	27.4407	27.4387	76.1408	35.6125	35.4875	35.5677
	CI Lower	23.5100	19.3260	19.1909	19.2736	45.5933	27.9283	27.3808	27.3763	75.6072	35.4995	35.3753	35.4579
	CI Upper	21.1170	19.3812	19.2987	19.2987	42.2428	27.9861	27.4982	27.5073	73.4682	35.7265	35.5792	35.8927
8	Avg	21.0205	19.3258	19.2441	19.2441	42.0472	27.9320	27.4385	27.4445	72.9893	35.6121	35.4673	35.7576
	CI Lower	20.9240	19.2703	19.1895	19.1895	41.8517	27.8779	27.3787	27.3817	72.5103	35.4977	35.3554	35.6224
	CI Upper	23.7587	19.4921	19.3323	19.3916	44.7553	27.8410	27.4989	27.4982	76.7684	35.6309	35.8124	37.3736
16	Avg	23.6740	19.4024	19.2854	19.3478	44.2760	27.7406	27.4321	27.4350	76.2435	35.5212	35.7242	36.9965
	CI Lower	23.5894	19.3128	19.2385	19.3039	43.7966	27.6401	27.3653	27.3717	75.7187	35.4115	35.6360	36.6194

Table A.11: 95% confidence intervals for the average costs, initial base-stock levels found via the disretized myopic policy vs FPA, for unknown transition matrix and demand distribution case, $N = 3$, $R_P = 7$, $c = 1$, $h = 1$ $b = 10$

n	95% CI	L = 0				L = 1				L = 2			
		Initial	R = 50	R = 200	R = 500	Initial	R = 50	R = 200	R = 500	Initial	R = 50	R = 200	R = 500
	CI Upper	16.2335	16.0792	16.1255	16.0565	22.3457	21.6625	21.6375	21.6054	28.3674	27.9931	27.9487	27.9652
4	Avg	16.1424	15.9993	16.0390	15.9759	22.2209	21.5754	21.5489	21.5150	28.2280	27.8735	27.8348	27.8532
	CI Lower	16.0512	15.9195	15.9524	15.8953	22.0961	21.4884	21.4604	21.4245	28.0886	27.7539	27.7210	27.7411
	CI Upper	16.4264	16.2519	16.3290	16.1383	22.1347	21.7433	21.6199	21.8720	29.2125	28.0505	27.9629	28.0080
8	Avg	16.2892	16.1777	16.2529	16.0555	22.0612	21.6536	21.5411	21.7917	28.8895	27.9576	27.8471	27.8959
	CI Lower	16.1520	16.1036	16.1768	15.9726	21.9877	21.5639	21.4624	21.7115	28.5666	27.8648	27.7313	27.7839
	CI Upper	16.0993	16.1940	16.1674	16.1372	21.9373	21.7281	21.7498	21.8780	28.1919	28.0359	27.9809	27.9619
16	Avg	16.0218	16.1088	16.0805	16.0588	21.8346	21.6317	21.6615	21.7754	28.1162	27.9230	27.8644	27.8554
	CI Lower	15.9444	16.0237	15.9935	15.9803	21.7319	21.5352	21.5733	21.6727	28.0405	27.8101	27.7480	27.7488

Table A.12: 95% confidence intervals for the average costs, initial base-stock levels found via the disretized myopic policy vs FPA, for unknown transition matrix and demand distribution case, $N = 4$, $R_P = 7$, $c = 1$, $h = 1$ $b = 10$

n	95% CI	L = 0				L = 1				L = 2			
		Initial	R = 50	R = 200	R = 500	Initial	R = 50	R = 200	R = 500	Initial	R = 50	R = 200	R = 500
	CI Upper	15.3565	15.4317	15.1607	15.1355	22.4659	19.9358	19.5563	19.4907	33.9265	25.3122	26.1079	26.9348
4	Avg	15.2226	15.3166	15.0387	15.0132	22.3594	19.8199	19.4389	19.3721	33.5940	25.2099	26.0082	26.8206
	CI Lower	15.0886	15.2015	14.9167	14.8908	22.2528	19.7040	19.3215	19.2534	33.2616	25.1076	25.9085	26.7064
	CI Upper	15.2781	15.2842	15.2323	15.1474	22.6065	19.8912	19.6040	19.6096	33.9782	25.7143	26.5446	27.5046
8	Avg	15.1589	15.1578	15.1012	15.0203	22.5024	19.7651	19.4805	19.4862	33.6620	25.6324	26.4601	27.4051
	CI Lower	15.0397	15.0314	14.9702	14.8933	22.3983	19.6390	19.3570	19.3627	33.3458	25.5505	26.3756	27.3057
	CI Upper	15.1514	15.1993	15.1589	15.1553	22.4992	20.0040	19.5551	19.4976	34.2079	26.8096	27.8798	29.2933
16	Avg	15.0333	15.0807	15.0389	15.0349	22.3969	19.8873	19.4411	19.3804	33.8807	26.6634	27.5974	28.8276
	CI Lower	14.9153	14.9622	14.9189	14.9145	22.2946	19.7706	19.3270	19.2632	33.5535	26.5172	27.3149	28.3619

Table A.13: 95% confidence intervals for the average costs, continuous myopic policy vs FPA, for unknown transition matrix and demand distribution case, $N = 2$, $R_P = 7$, $c = 1$, $h = 1$ $b = 10$

n	95% CI	L = 0				L = 1				L = 2			
		Myopic	R = 50	R = 200	R = 500	Myopic	R = 50	R = 200	R = 500	Myopic	R = 50	R = 200	R = 500
	CI Upper	23.7921	19.4650	19.3017	19.3846	44.2985	28.0419	27.5006	27.5011	73.0556	35.7256	35.5998	35.6774
4	Avg	23.6944	19.3955	19.2463	19.3291	43.9245	27.9851	27.4407	27.4387	72.5777	35.6125	35.4875	35.5677
	CI Lower	23.5967	19.3260	19.1909	19.2736	43.5506	27.9283	27.3808	27.3763	72.0998	35.4995	35.3753	35.4579
	CI Upper	23.7921	19.3812	19.2987	19.2987	44.2985	27.9861	27.4982	27.5073	73.0556	35.7265	35.5792	35.8927
8	Avg	23.6944	19.3258	19.2441	19.2441	43.9245	27.9320	27.4385	27.4445	72.5777	35.6121	35.4673	35.7576
	CI Lower	23.5967	19.2703	19.1895	19.1895	43.5506	27.8779	27.3787	27.3817	72.0998	35.4977	35.3554	35.6224
	CI Upper	23.7921	19.4921	19.3323	19.3916	44.2985	27.8410	27.4989	27.4982	73.0556	35.6309	35.8124	37.3736
16	Avg	23.6944	19.4024	19.2854	19.3478	43.9245	27.7406	27.4321	27.4350	72.5777	35.5212	35.7242	36.9965
	CI Lower	23.5967	19.3128	19.2385	19.3039	43.5506	27.6401	27.3653	27.3717	72.0998	35.4115	35.6360	36.6194

Table A.14: 95% confidence intervals for the average costs, continuous myopic policy vs FPA, for unknown transition matrix and demand distribution case, $N = 3$, $R_P = 7$, $c = 1$, $h = 1$ $b = 10$

n	95% CI	L = 0				L = 1				L = 2			
		Myopic	R = 50	R = 200	R = 500	Myopic	R = 50	R = 200	R = 500	Myopic	R = 50	R = 200	R = 500
	CI Upper	16.1415	16.0792	16.1255	16.0565	25.4319	21.6625	21.6375	21.6054	40.6756	27.9931	27.9487	27.9652
4	Avg	16.0620	15.9993	16.0390	15.9759	25.3074	21.5754	21.5489	21.5150	40.3190	27.8735	27.8348	27.8532
	CI Lower	15.9825	15.9195	15.9524	15.8953	25.1829	21.4884	21.4604	21.4245	39.9624	27.7539	27.7210	27.7411
	CI Upper	16.1415	16.2519	16.3290	16.1383	25.4319	21.7433	21.6199	21.8720	40.6756	28.0505	27.9629	28.0080
8	Avg	16.0620	16.1777	16.2529	16.0555	25.3074	21.6536	21.5411	21.7917	40.3190	27.9576	27.8471	27.8959
	CI Lower	15.9825	16.1036	16.1768	15.9726	25.1829	21.5639	21.4624	21.7115	39.9624	27.8648	27.7313	27.7839
	CI Upper	16.1415	16.1940	16.1674	16.1372	25.4319	21.7281	21.7498	21.8780	40.6756	28.0359	27.9809	27.9619
16	Avg	16.0620	16.1088	16.0805	16.0588	25.3074	21.6317	21.6615	21.7754	40.3190	27.9230	27.8644	27.8554
	CI Lower	15.9825	16.0237	15.9935	15.9803	25.1829	21.5352	21.5733	21.6727	39.9624	27.8101	27.7480	27.7488

Table A.15: 95% confidence intervals for the average costs, continuous myopic policy vs FPA, for unknown transition matrix and demand distribution case, $N = 4$, $R_P = 7$, $c = 1$, $h = 1$ $b = 10$

n	95% CI	L = 0				L = 1				L = 2			
		Myopic	R = 50	R = 200	R = 500	Myopic	R = 50	R = 200	R = 500	Myopic	R = 50	R = 200	R = 500
	CI Upper	15.1072	15.4317	15.1607	15.1355	22.0328	19.9358	19.5563	19.4907	32.7779	25.3122	26.1079	26.9348
4	Avg	14.9878	15.3166	15.0387	15.0132	21.9362	19.8199	19.4389	19.3721	32.4912	25.2099	26.0082	26.8206
	CI Lower	14.8685	15.2015	14.9167	14.8908	21.8395	19.7040	19.3215	19.2534	32.2046	25.1076	25.9085	26.7064
	CI Upper	15.1072	15.2842	15.2323	15.1474	22.0328	19.8912	19.6040	19.6096	32.7779	25.7143	26.5446	27.5046
8	Avg	14.9878	15.1578	15.1012	15.0203	21.9362	19.7651	19.4805	19.4862	32.4912	25.6324	26.4601	27.4051
	CI Lower	14.8685	15.0314	14.9702	14.8933	21.8395	19.6390	19.3570	19.3627	32.2046	25.5505	26.3756	27.3057
	CI Upper	15.1072	15.1993	15.1589	15.1553	22.0328	20.0040	19.5551	19.4976	32.7779	26.8096	27.8798	29.2933
16	Avg	14.9878	15.0807	15.0389	15.0349	21.9362	19.8873	19.4411	19.3804	32.4912	26.6634	27.5974	28.8276
	CI Lower	14.8685	14.9622	14.9189	14.9145	21.8395	19.7706	19.3270	19.2632	32.2046	26.5172	27.3149	28.3619